



**HAL**  
open science

# Simplicité et complexité des langues dans l'histoire des théories linguistiques

Chloé Laplantine, John E Joseph, Emilie Aussant

► **To cite this version:**

Chloé Laplantine, John E Joseph, Emilie Aussant (Dir.). Simplicité et complexité des langues dans l'histoire des théories linguistiques. SHESL. SHESL, 2023, HEL Livres, Chloé Laplantine, 979-10-91587-21-1. hal-04147802

**HAL Id: hal-04147802**

**<https://hal.science/hal-04147802>**

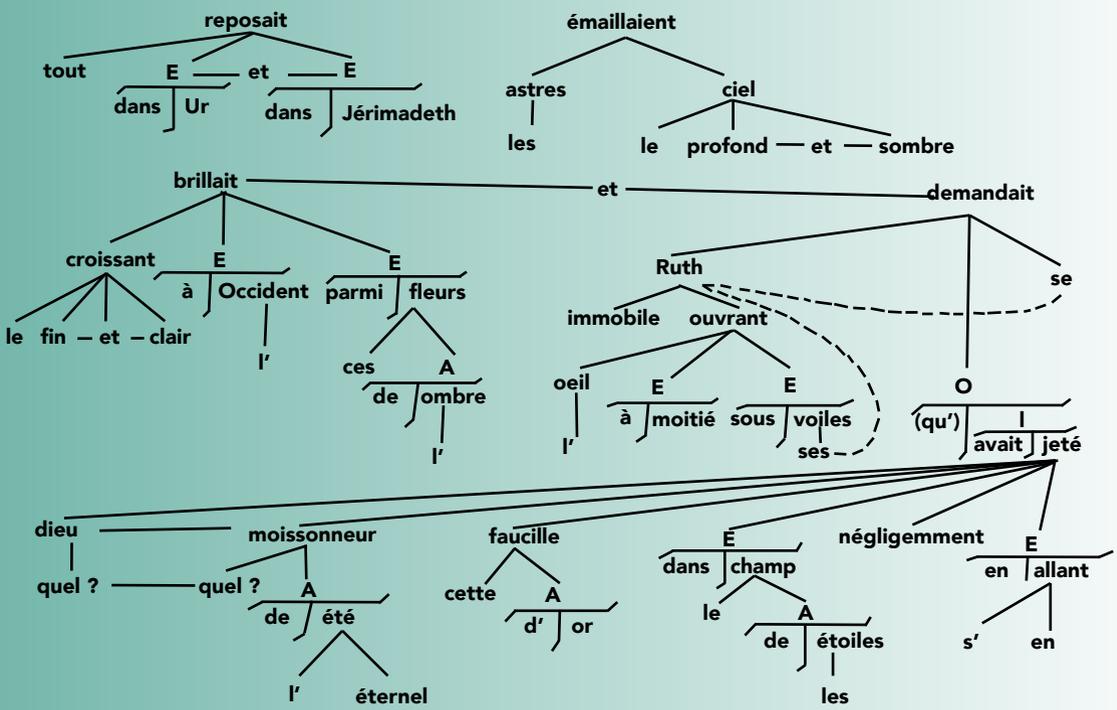
Submitted on 1 Jul 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



# Simplicité et complexité des langues dans l'histoire des théories linguistiques

Sous la direction de  
Chloé Laplantine, John E. Joseph et Émilie Aussant

**SHESL**

Simplicité et complexité des langues  
dans l'histoire des théories  
linguistiques



# Simplicité et complexité des langues dans l'histoire des théories linguistiques

Actes du colloque de la  
Société d'histoire et d'épistémologie des sciences du langage  
et du laboratoire Histoire des théories linguistiques  
Paris, 23-25 janvier 2020

Sous la direction de  
Chloé Laplantine, John E. Joseph & Émilie Aussant

**SHESL**

HEL Livres

HEL Livres est une collection spécialisée dans le domaine de l'histoire et de l'épistémologie de la linguistique.

Elle accueille les projets monographiques ou collectifs, les éditions et les traductions de textes relevant de ce domaine de recherche.

Elle est dotée d'un comité éditorial identique au comité de rédaction de la revue *Histoire Épistémologie Langage (HEL)*.

Les ouvrages de la collection sont disponibles en libre accès au format numérique et peuvent être acquis au format imprimé.

Titres parus :

Samain, Didier & Pierre-Yves Testenoire, dir. 2022. *La linguistique et ses formes historiques d'organisation et de production*. Paris : SHESL (HEL Livres, 1).

Candel, Danielle, Didier Samain & Dan Savatovsky, dir. 2022. *Eugen Wüster et la terminologie de l'école de Vienne*. Paris : SHESL (HEL Livres, 2).

Chloé Laplantine, John E. Joseph & Émilie Aussant, dir. 2023. *Simplicité et complexité des langues dans l'histoire des théories linguistiques*. Paris : SHESL (HEL Livres, 3).

Cet ouvrage est paru avec le soutien de l'Université Paris Cité.

Les textes de cette publication relèvent de la licence Creative Commons CC BY-NC-ND 4.0 : partage autorisé sous condition d'attribution des œuvres à leurs auteurs, réutilisation commerciale et modification interdites.

 <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.fr>

ISBN : 979-10-91587-21-1

Collection HEL Livres

Directrice : Chloé Laplantine

ISSN 2968-5222 (imprimé)

ISSN 2967-3321 (en ligne)

Éditeur (secrétariat d'édition) : Chloé Laplantine

Conception maquette intérieure : Virginie Teillet/Italiques

Couverture : Chloé Laplantine

SHESL – Société d'histoire et d'épistémologie des sciences du langage

Université Paris Cité – Case 7034

27 rue Jean-Antoine-de-Baïf

75 013 Paris (France)

<https://shesl.org>

# Les auteurs

**Marie-Madeleine Bertucci** est professeure de sciences du langage à CY Cergy Paris université. Sociolinguiste, elle s'intéresse aux normes endogènes de français et a dirigé à ce sujet l'ouvrage collectif *Les français régionaux dans l'espace francophone* (Peter Lang, 2016) ainsi qu'aux pratiques langagières des jeunes dans les contextes urbains ségrégués plurilingues à propos desquelles elle a publié plusieurs articles dans lesquels elle aborde aussi les problèmes relatifs à l'appropriation-apprentissage par ces derniers de la langue de scolarisation. Elle s'intéresse également à des questions de didactique et a publié en 2020 un ouvrage intitulé *Étude grammaticale et stylistique de textes de langue française. CAPES et agrégation de lettres modernes* (université Grenoble-Alpes éditions).

**Maria Chriti** is Teaching Assistant of Classics at Thessaloniki's Aristotle University. She holds a degree, an MA and a Ph.D. in Classics from the same University. She has been a Post-doctoral fellow at the Centre for Hellenic Studies / University of Harvard (2017-2018). She was a research associate in the Divisions of Linguistics and Lexicography (1999-2018) at the Centre of the Greek Language (Greek Ministry of Education). She also taught ancient Greek philosophy at the Open University of Cyprus (2016-2018), while she has been the President of the International Association "ETYGRAM" (*Etymologies Grecques Antiques et Médiévales*; <http://www.cepam.cnrs.fr/etygram/>). Her research interests include ancient Greek linguistics/philosophy of language, Neoplatonism, ancient Greek etymology. Her publications include *A History of Ancient Greek: From the Beginnings to Late Antiquity* (in collaboration with A.-F. Christidis & M. Arapopoulou, Cambridge: CUP 2007), "Ancient Philosophers on Language" (in collaboration with P. Kotzia [†], in *The Encyclopedia of Ancient Greek Language and Linguistics* (eds. G. Giannakis et al. vol. I, 124-133. Leiden, Boston: Brill 2013).

**José Edicarlos de Aquino** est titulaire d'une thèse en Sciences du langage par l'Universidade Estadual de Campinas en co-tutelle avec l'Université Sorbonne Nouvelle. Il est le coordinateur du Laboratoire d'études linguistiques à l'Universidade Federal do Tocantins. Il développe

des recherches dans le domaine de l'histoire des idées linguistiques et étudie la réflexion sur le langage et les langues au XIX<sup>e</sup> siècle au Brésil, principalement la question de la production grammaticale et les relations entre littérature et outillage linguistique.

**Jacques François** est professeur émérite à l'université de Caen-Normandie, après avoir exercé aux universités de Constance et de Nancy. Il a publié plusieurs monographies sur la syntaxe et sémantique du français et de l'allemand et il conduit actuellement un projet de modélisation graphique de la polysémie évolutive sur la base des notices historiques du Trésor de la Langue Française. En historiographie de la linguistique, il est notamment l'auteur du *Siècle d'or de la linguistique en Allemagne - de Humboldt à Meyer-Lübke* (Lambert-Lucas, 2018) et de *Johann-Christoph Adelung, linguiste des Lumières à la cour de Saxe* (L'Harmattan, 2020). Il s'est vu attribuer le prix Émile Benveniste de l'Académie des Inscriptions et Belles Lettres en 2023.

**Jean-Michel Fortis** est chargé de recherche au CNRS. Il est membre de l'équipe Histoire des théories linguistiques (Université Paris Cité & Université Sorbonne Nouvelle). Ses travaux sont consacrés principalement à l'histoire de la linguistique américaine récente, aux rapports entre psychologie et linguistique, à l'expression des relations spatiales et à la question de leur rôle au sein de la théorie linguistique (localisme). Il codirige, avec Otto Zwartjes, la revue *Historiographia Linguistica*.

**Nathalie Fournier**, professeur émérite à l'université Lumière-Lyon2, est spécialiste d'histoire de la langue et de la grammaire du français classique. Elle est l'auteur d'une *Grammaire du français classique* (1998) et a publié récemment une édition critique de la *Grammaire et syntaxe française* de Charles Maupas (1618).

**Gerda Haßler** est présidente de la Société Leibniz des Arts et Sciences à Berlin et professeur émérite de linguistique théorique et appliquée de l'Université de Potsdam. Elle a soutenu sa thèse en 1978 à l'Université de Halle où elle a présenté sa thèse d'Habilitation en 1984. Ses domaines de recherche sont l'histoire de la linguistique, la syntaxe et la pragmatique des langues romanes et l'étude comparée de la lexicologie. Elle a publié, entre autres, les livres suivants : *Lexikon sprachtheoretischer Grundbegriffe des 17. und 18. Jahrhunderts* (avec Cordula Neis, De

Gruyter 2009), *Temporalität, Aspektualität und Modalität in romanischen Sprachen* (De Gruyter 2016), *Collocations et traditions discursives* (LINX 2020), *Brevitas et prolixitas* (avec Friederike Spitzl-Dupic, Nodus 2021), *Manuel des modes et modalités* (De Gruyter 2022).

**Sophie Jollin-Bertocchi** est Maître de conférences habilitée à diriger des recherches à l'Université Paris-Saclay, UVSQ, et membre de l'EA Centre d'Histoire Culturelle des Sociétés Contemporaines. Ses activités de recherche portent sur plusieurs domaines : les représentations de la langue chez les écrivains, la linguistique textuelle, l'histoire des idées esthétiques et linguistiques. Co-éd. avec Serge Linarès *Changer de style. Écritures évolutives aux XX<sup>e</sup> et XXI<sup>e</sup> siècles*, Brill, Rodopi, 2021.

**John E. Joseph** est Professeur de Linguistique Appliquée à l'Université d'Édimbourg. Il est l'auteur d'*Eloquence and Power : The Rise of Language Standards and Standard Languages* (Blackwell, 1987), *Limiting the Arbitrary: Naturalism and its Opposites in Plato's Cratylus and Modern Theories of Language* (Benjamins, 2000), *From Whitney to Chomsky: Essays in the History of American Linguistics* (Benjamins, 2002), *Language and Identity: National, Ethnic, Religious* (Palgrave Macmillan, 2004), *Language and Politics* (Edinburgh University Press, 2006), *Saussure* (Oxford University Press, 2012 ; version française, traduite par Nathalie Vincent-Arnaud, Lambert-Lucas, 2021) et *Language, Mind and Body: A Conceptual History* (Cambridge University Press, 2018). Il a traduit les *Dernières leçons* d'Émile Benveniste (*Last lectures : Collège de France, 1968 and 1969*, Edinburgh University Press, 2019). L'Académie royale danoise des sciences et des lettres l'a élu comme membre correspondant. Il siège aux comités de plusieurs revues, séries et sociétés, y compris le Cercle Ferdinand de Saussure, dont il est actuellement président.

**Chloé Laplantine** est chargée de recherche au CNRS, membre du laboratoire Histoire des théories linguistiques. Elle a publié différents travaux sur l'articulation de la linguistique et de la poétique chez Émile Benveniste, Edward Sapir, et Ferdinand de Saussure. Elle s'intéresse à l'histoire de l'approche des langues et cultures autochtones d'Amérique du Nord aux 19<sup>e</sup> et 20<sup>e</sup> siècles. Elle a consacré plusieurs études à la linguistique impulsée par Franz Boas, ainsi qu'à l'importance des textes dans l'invention de l'ethnolinguistique nord-américaine. Elle prépare une édition critique des carnets de terrain d'Émile Benveniste sur plusieurs langues de la côte Pacifique Nord-Ouest.

**Nikolaos Lavidas** is Associate Professor of Diachronic Linguistics at the Faculty of English Language and Literature (Department of Language and Linguistics), School of Philosophy, National and Kapodistrian University of Athens. His research interests lie in the areas of language change, (historical) language contact, historical corpora, and syntax-semantics interface.

**Maxime Maleux** studied Classics at KU Leuven. After he obtained his MA in 2017, he studied Hebrew at the same institution. In October 2018, he started working on the Hebrew part of the project ‘*Ad fontes!*’ in the Classroom: Teaching Latin, Greek, and Hebrew Texts in the Early Modern Southern Low Countries, under supervision of Prof. dr. Pierre Van Hecke and dr. Raf Van Rooy. In November 2020, he was granted a PhD fellowship by Research Foundation - Flanders (FWO). The aim of his research is to clarify how Hebrew was taught in the sixteenth-century *Collegium Trilingue Lovaniense* (the *Three Languages College* in Leuven) by investigating Hebrew grammars written and used in the Low Countries, as well as course notes made by students of the Trilingue. One of his publications related to this topic is the article “On Hebrew Nymphs and Aqueducts. Two Hebrew Poems by Andreas Masius”, published in *Lias: Journal of Early Modern Intellectual Culture and Its Sources* (2020).

**Cendrine Pagani-Naudet** est agrégée de Lettres modernes et enseigne dans le secondaire. Ses travaux portent sur l’histoire de la langue et de la grammaire française. Elle est l’auteur d’un ouvrage sur la dislocation (*La dislocation xix<sup>e</sup>-xvii<sup>e</sup>. Histoire d’un procédé de style*) et l’édition critique de l’*Essay d’une parfaite grammaire de la langue française* de Laurent Chiflet (1659).

**Mat Pires** is a lecturer in the English department of the University of Franche-Comté, Besançon, specializing in sociolinguistics and discourse analysis. He has published on the discourse of music journalism (subject of a PhD at the University of Surrey), neology in militant contexts and in the work of Lewis Carroll, and conflict between linguistic norms and usage. He is the English translator of Saussure’s *Écrits de linguistique générale* (*Writings in General Linguistics*, OUP).

**Jacques-Philippe Saint-Gerand**, retraité, Paris, Poitiers, Clermont-Ferrand, Limoges, a enseigné la linguistique au siècle dernier, selon trois

axes principaux : Histoire des idées linguistiques, Histoire de la langue et Études de style. *Mutations et sclérose : la langue française 1789-1848* (éd.), Franz Steiner Verlag, 1993, Stuttgart. *Langue et Linguistique : Mouvements croisés et alternés (1790-1860)*; *Aspects involutifs de l'historiographie linguistique française au XIX<sup>e</sup> siècle*, *Travaux de Linguistique* 33, Louvain, 1997.

**Claudia Schweitzer** est titulaire de deux thèses en Sciences du langage et en musicologie. Cette double formation lui permet des approches interdisciplinaires dans ses recherches qui s'intéressent le plus souvent à l'âge classique français. En 2022, elle a publié deux ouvrages intitulés *Die Musik der Sprache. Französische Prosodie im Spiegel der musikalischen Entwicklungen vom 16. bis 21. Jahrhundert* (Narr) et *La voix, souffle de l'émotion. Parole et chant à l'âge classique* (ENS).

**Julien Sibileau** est Maître de Conférences en linguistique et histoire de la grammaire arabe à l'Institut National des Langues et Civilisations Orientales (INaLCO) et membre du Centre de Recherches Moyen-Orient et Méditerranée (CERMOM). Ses spécialités sont la linguistique arabe et l'histoire de la grammaire arabe. Il est l'auteur d'une thèse intitulée : *Une entreprise de légitimation de la grammaire arabe : les Ḥaṣā'is d'Ibn Ğinnī*, soutenue en 2021.

**Vassilis Symeonidis** is PhD Candidate of Linguistics at the Faculty of English Language and Literature (Department of Language and Linguistics), School of Philosophy, National and Kapodistrian University of Athens. His research interests lie at the interface between theoretical and quantitative linguistics with a particular focus on syntax, historical linguistics and language change.

**Margaret Thomas** is Professor of Linguistics at Boston College, in Chestnut Hill, Massachusetts USA. Her research interests are broadly in the history of linguistics, especially twentieth-century linguistics in the United States. Her most recent book is *Formalism and Functionalism in Linguistics: The Engineer and the Collector* (Routledge, 2020). She is now at work on a monograph about the history of racism in American linguistics, up to the early years of the Linguistic Society of America.

**Raf Van Rooy** (PhD in Linguistics, KU Leuven, 2017) is assistant professor of Latin Literature at KU Leuven. Van Rooy previously was PhD fellow (2013–17) and postdoctoral fellow (2017–21) of the Research

Foundation–Flanders (FWO) at KU Leuven, and MSCA postdoctoral researcher at the University of Oslo (2021–22). He holds degrees in Classics, General Linguistics, and Early Modern History from KU Leuven, UCLouvain, and Ghent University, and specializes in the premodern history of linguistics and Greek studies in the Neo-Latin world. From 2020 until 2023, he was editor at the peer-reviewed journal *Language & History*. Van Rooy is the author of *Language or Dialect?* (Oxford University Press), *Greece’s Labyrinth of Language* (Language Science Press), and *New Ancient Greek in a Neo-Latin World* (Brill). He is currently also supervising the development of DaLeT, the Database of the Leuven Trilingue, and co-editor of *Humanistica Lovaniensia*.

**Alice Vittrant** est Professeur de linguistique à l’Université Lumière Lyon 2 et membre du laboratoire Dynamique Du Langage (CNRS-DDL). Ses spécialités sont la typologie, l’expression spatiale, le contact des langues et les langues d’Asie du Sud-Est – en particulier le birman et le hmong. Elle est l’auteur de plusieurs articles et chapitres de livres sur des aspects de la grammaire birmane. Elle a également coédité un ouvrage de référence sur la linguistique des langues de l’Asie du Sud-Est continentale (Vittrant & Watkins 2019).

# Introduction

John E. Joseph<sup>1</sup> et Chloé Laplantine<sup>2</sup>

<sup>1</sup> Université d'Édimbourg

<sup>2</sup> Université Paris Cité and Université Sorbonne Nouvelle, CNRS, Laboratoire d'histoire des théories linguistiques, F-75013 Paris, France

En tant que linguistes, il nous est interdit en principe de formuler des jugements de qualités sur une langue ou un dialecte, ou même sur une forme ou une structure linguistique. Un moment important dans le développement de cette interdiction est l'échange entre Wilhelm von Humboldt et certains de ses lecteurs, qui avaient vu dans ses écrits des préjugés en faveur des langues indo-européennes, surtout le grec ancien et le sanskrit, ceci impliquant en conséquence une infériorité des autres familles linguistiques.

Aux États-Unis, John Pickering et Peter Stephen Du Ponceau protestèrent en faveur des langues américaines, et en France, le sinologue Jean-Pierre Abel-Rémusat fit de même pour le chinois ancien. Dans une longue lettre adressée à Abel-Rémusat, et publiée en 1826 et 1827, Humboldt expliquait que la structure des langues flexionnelles telles que le sanskrit, et des langues isolantes telles que le chinois, les rend chacune supérieure pour des fonctions mentales différentes : le chinois pour l'expression des *idées* prises singulièrement, le sanscrit pour celle de la *pensée* dans son intégralité.

Cette distinction est d'un ordre *simple-complexe*. Elle fournit aux linguistes le moyen d'avoir le beurre et l'argent du beurre : en localisant la supériorité des langues flexionnelles dans leur capacité à combiner les idées – même si l'expression des idées individuelles est meilleure dans les langues isolantes – on ne devait pas se débarrasser de la perception occidentale générale d'un Orient relativement

primitif. En même temps, on était libre de proclamer la supériorité du primitif, comme dans le trope, d'inspiration rousseauiste, du Bon Sauvage.

Bien que ce trope soit plus caractéristique de la littérature et des arts graphiques que des sciences, Arthur Lovejoy et George Boas (1935 : 7) ont identifié un « primitivisme culturel », de date ancienne, qui consiste dans « la conviction des hommes vivant dans une condition culturelle relativement évoluée et complexe qu'une vie beaucoup plus simple et moins sophistiquée, à certains égards ou à tous égards, est une vie plus désirable »<sup>1</sup>. Plus désirable peut-être que la vie complexe de la civilisation, et plus aisée à certains égards, mais pas à tous les égards :

Elle est plus facile précisément *parce que* c'est (ou l'on a imaginé, en partie à tort, qu'elle était) plus simple ; elle est moins chargée de dispositifs et (comme on l'a supposé) d'une multitude de règles, de règlements et de conventions restrictifs. [...] D'autre part, la vie de nombreux peuples sauvages, et ceux qui sont les mieux connus dans la plus grande partie de l'histoire européenne, est manifestement à certains égards, ou lorsqu'elle est considérée sous un certain jour, plus dure que celle de l'humanité civilisée – au moins de la partie la plus prospère. (*Ibid.* : 9-10 ; notre traduction)<sup>2</sup>

Ce paradoxe suffit pour autoriser une disjonction entre le primitivisme et la hiérarchisation qualitative des peuples et des langues. Mais pour les linguistes, il était évident que l'histoire n'est pas un simple processus de complexification ; au contraire, dans la famille linguistique la mieux connue, l'indo-européen, le système des flexions des langues anciennes était bien plus complexe que dans les langues modernes, même dans le lituanien, qui conserve un nombre

---

1. « ... *the belief of men living in a relatively highly evolved and complex cultural condition that a life far simpler and less sophisticated in some or in all respects is a more desirable life* » (1965 :7).

2. « *It is easier precisely because it is (or has, in part erroneously, been imagined to be) simpler; it is less burdened with apparatus and (as has been supposed) with a multitude of restrictive rules and regulations and conventionalities. [...] On the other hand, the life of many savage peoples, and those the best known throughout the greater part of European history, is manifestly in certain respects, or when regarded in a certain light, harder than that of civilized mankind – at least of the more prosperous portion of it* » (1965 : 9-10).

exceptionnel de formes. La morphologie n'est pas tout le système linguistique, bien sûr, et la phonologie comme la syntaxe ont subi des développements divers. Au fur et à mesure que la linguistique devenait une « science », les linguistes cessèrent de spéculer sur l'état de la langue à son aube ; et pour les modernes, la science commença au moment où la comparaison des langues-filles attestées permit de reconstruire la langue-mère d'une famille quelconque.

L'origine des langues n'est ressuscitée comme sujet scientifique autorisé que dans les années 1980, avec une présomption fondamentale inchangée, à savoir que les langues sont passées d'une origine simple à un stade complexe, qui, dans certains cas, y compris celui des langues indo-européennes, a subi ensuite une simplification, ponctuée par des complexifications qui restaient plutôt exceptionnelles. Mais, entre-temps, des événements historiques avaient périodiquement modifié l'atmosphère anthropologique et politique. Dès 1850, Arthur de Gobineau offrait un « racisme scientifique », en fait dépourvu de valeur scientifique mais réfléchissant les préjugés reçus ; et neuf ans plus tard, l'évolutionnisme de Charles Darwin établissait l'image d'une échelle de l'évolution commençant avec les êtres vivants les plus simples et continuant jusqu'aux êtres humains. Bien que vraiment scientifique, le darwinisme parut soutenir, pour ceux qui désiraient une telle confirmation, l'illusion d'une hiérarchie des races humaines.

Les prolongements de cette manière inhumaine de concevoir l'humanité sont trop bien connus. En linguistique, ils existèrent, mais marginalement : Léopold de Saussure, par exemple, proclama une échelle de qualité linguistique liée à une échelle de complexité, dans des écrits lus par peu de gens ; par contre, son frère Ferdinand s'éloigna des comparaisons de ce genre, en faveur d'une conception du système linguistique comme un réseau de pures différences, et nia la possibilité d'une liaison entre langue et culture, dans ses cours qui ont trouvé des lecteurs innombrables. Aux États-Unis, Franz Boas, bien que dévoué à l'idée d'un nœud étroit entre langue et culture, affirma robustement l'égalité des langues et des peuples généralement divisés entre primitifs et civilisés.

Mais les perceptions fortes ne peuvent jamais être anéanties pour toujours. L'éternel retour nietzschéen de l'analyse des langues selon

leur simplicité et complexité recommencera au début des années 1930, dans les ouvrages de Roman Jakobson et Nikolai Troubetzkoy, lorsqu'ils formuleront ensemble la notion de « marque ». Là où, dans le système saussurien, chaque élément a un statut égal en tant que différence, Jakobson et Troubetzkoy affirment au contraire que certains éléments sont plus simples que d'autres, et que les éléments simples sont distingués des éléments complexes par l'addition d'une marque supplémentaire. Troubetzkoy mourut peu après l'invasion allemande de l'Autriche en 1938, mais Jakobson continua le développement de ces idées, proposant que les éléments simples sont acquis plus tôt par les enfants, sont plus durables à travers l'histoire de la langue, et ainsi de suite. Il dut fuir les Nazis, et fit un passage vers la Suède, où il put publier un livre dont l'importance fut reconnue immédiatement mais dont l'influence fut retardée par la guerre.

À la fin des hostilités, le choc de la découverte du génocide allemand eut un effet profond à travers le monde. Une analyse basée sur la simplicité et la complexité n'attirait pas grand nombre de linguistes, jusque dans les années 1960, et ce n'est qu'assez récemment qu'a reculé le dogme reçu formulé dans l'après-guerre que « toutes les langues sont également complexes » (voir Joseph & Newmeyer 2012). À cause de ce dogme, des recherches sur la simplicité et la complexité des langues sont restées effectivement suspectes jusqu'aux années 1980 – et même à cette époque-là, les linguistes assez courageux pour les ressusciter il y a quarante ans reçurent parfois des accusations de vouloir rouvrir le chemin menant à Auschwitz. Peu à peu, ce champ de recherches a retrouvé la reconnaissance de sa validité et de son importance pour l'explication de bien des phénomènes linguistiques, diachroniques et synchroniques, théoriques et appliqués.

Les articles ici rassemblés contribuent à notre connaissance de ce trajet historique dans la pensée linguistique, et de l'applicabilité des mesures de complexité dans la linguistique actuelle. Nous, les contributeurs, sommes tous reconnaissants des efforts de nos prédécesseurs qui ont rendu possible la poursuite des intérêts scientifiques qui a abouti à l'ouvrage que voici.

Cet ouvrage constitue les actes du colloque « Simplicité et complexité des langues dans l'histoire des théories linguistiques » organisé par la Société d'histoire épistémologie

des sciences du langage (SHESL), qui s'est tenu à Paris du 23 au 25 janvier 2020.

Il nous a semblé que les textes rassemblés ici pouvaient être regroupés selon cinq fils directeurs.

Un premier fil est celui *du simple, du naturel et du natif*. Gerda Hassler examine la réflexion sur l'ordre fixe des mots comme critère d'évaluation des langues dans les grammaires du <sup>xvi</sup>e au <sup>xviii</sup>e siècles, et met en lumière la réapparition de ce critère d'évaluation dans les théories modernes de la structure informationnelle où l'ordre fixe ou libre des mots est conçu en termes d'effort de traitement et de décodage. Hassler rappelle que l'ordre naturel est discuté dès l'Antiquité (par exemple chez Denys d'Halicarnasse, Quintilien, Cicéron). Notion de rhétorique, elle est transférée à la grammaire, et devient un critère prescriptif notamment chez Louis Meigret (« *construccion ou ordonance bone de parolles* ») où le français prend valeur d'exemple par rapport au latin. Chez Arnauld et Lancelot, s'il existe une *construction simple* ou *régulière* qui correspond à un ordre naturel de l'expression des pensées, l'usage d'une *construction irrégulière* ou *figurée* est également légitime dans un but de brièveté et d'élégance. *Clarté, netteté, naïveté* du discours : on voit se poursuivre cette réflexion portée par des idées rationalistes jusqu'au développement de la pensée sensualiste au <sup>xviii</sup>e siècle : Condillac pose que le « naturel varie nécessairement selon le génie de la langue », Batteux contredit le dogme rationaliste en voyant dans le français une langue inversive, et Diderot propose l'étude comparée de la langue des gestes, comme premier stade du langage, et du langage articulé. Dans le domaine de la normalisation des langues, on voit le modèle rationaliste dominer chez les grammairiens français du <sup>xviii</sup>e siècle, qui en même temps posent la langue française en exemple d'une langue suivant l'ordre naturel des idées. Hassler examine par la suite la problématique de l'ordre des mots telle qu'elle a été discutée en Europe (Espagne, Allemagne, Russie, Angleterre). Par exemple, en Espagne, chez Nebrija, l'ordre naturel ne doit pas aller contre l'ordre du réel, principe que suit la Real Academia Española ; les grammairiens espagnols semblent promouvoir une liberté de l'ordre des mots français pour des fins stylistiques (élégance et énergie) ou

communicationnelles. La discussion sur l'ordre naturel des langues, qu'on pourrait croire confinée à l'époque classique, se poursuit en fait encore chez des auteurs comme Henri Weill, Georg von der Gabelentz et Hermann Paul au XIX<sup>e</sup> siècle. Le grammairien Wilhelm Meyer-Lübke parle ainsi de lutte entre des facteurs grammaticaux et affectifs. Et comme on l'a dit plus haut, cette discussion a toute son actualité aujourd'hui.

Claudia Schweitzer, dans « Pourquoi l'italien serait-il la « meilleure » langue pour chanter ? ou : Comment la sonorité peut déterminer l'impression de simplicité d'une langue », montre que les conceptions qui touchent à l'évaluation des qualités des langues dans leur rapport au chant ont varié dans l'histoire. L'opinion populaire, qui inclut celle des chanteurs eux-mêmes, semble s'accorder à penser que l'italien est la meilleure langue pour chanter, et on voit que cette opinion est déjà établie au XVIII<sup>e</sup> siècle, par exemple dans les écrits de Jean-Jacques Rousseau qui trouve que le français n'a ni mesure ni mélodie. A l'âge baroque, une union étroite est conçue entre langue et musique : le texte n'est pas posé sur la musique, mais est la source même du chant, et la reconnaissance de qualités propres aux langues aboutit à la conception de styles différents qui leur correspondent. L'italien est néanmoins reconnu naturellement plus expressif de par ses sonorités vocaliques, et donc naturellement plus adapté au chant, ces représentations de la langue ayant des incidences sur les techniques de chants en italien et dans les autres langues.

Sophie Jollin-Bertocchi et Jacques-Philippe Saint-Gerand mènent l'enquête sur le couple terminologique phrase simple / phrase complexe depuis le XVII<sup>e</sup> siècle jusqu'aux grammaires contemporaines, et mettent ainsi en lumière différentes étapes de la réflexion dans l'approche de la notion de phrase. Dans la plupart des grammaires de la fin du XX<sup>e</sup> siècle, le couple phrase simple / complexe est présenté comme traditionnel, alors que son officialisation est récente, les grammaires du début du siècle marquant rarement cette distinction. La distinction entre phrase simple et phrase complexe apparaît d'abord dans une tension entre les raisonnements de la grammaire et de la rhétorique ; ainsi une proposition simple peut être déployée en une période oratoire complexe. Dans la grammaire de Port-Royal, une phrase peut être analysée en propositions principale et incidente,

chacune étant l'expression d'un jugement. A la fin du XVIII<sup>e</sup> siècle et dans la première partie du XIX<sup>e</sup> siècle, c'est de nouveau la rhétoricité de la langue qui détermine l'approche de la phrase, avec, par exemple chez Napoléon Landais, la différenciation entre *construction analytique* (où la construction de la phrase suit l'ordre naturel des idées) et *construction figurée* (qui suit l'ordre propre des langues). Dans la seconde moitié du XIX<sup>e</sup> siècle, on voit se développer l'analyse logique de la phrase en propositions, qui n'a plus rien à voir avec la proposition dans son sens « métaphysique ». A partir de la fin du XIX<sup>e</sup> siècle, la grammaire scolaire développe des stratégies différentes pour des audiences différentes : l'enseignement grammatical de la phrase simple est destiné à l'enseignement primaire et celui de la phrase complexe, plus rhétorique, destiné à l'enseignement secondaire des classes sociales favorisées. Le couple se maintient et est même structurant dans les grammaires à visée didactique du XX<sup>e</sup> siècle et jusqu'à maintenant, aux côtés d'autres termes comme phrase *matrice*, *constituante*, *multiple*, etc., qui témoignent d'une difficulté à opérer seulement avec la distinction simple / complexe, dont les frontières fragiles peuvent être marquées par la subordination, la coordination ou la juxtaposition, ou encore par les phénomènes de jonction ou de translation (chez Lucien Tesnière). Finalement, il apparaît que le couple phrase simple / complexe porte le fardeau d'une histoire indécise et tiraillée entre grammaire et rhétorique.

Cendrine Pagani-Naudet et Nathalie Fournier s'intéressent à la notion de *naïf* appliquée à la langue, qui apparaît à la Renaissance comme un double de *natif*, mais avec des valeurs différentes, et qui bougent entre XVI<sup>e</sup> et XVII<sup>e</sup> siècles selon l'idéologie linguistique qui en guide l'approche. Ainsi sont dégagées trois valeurs différentes et qui vont s'affronter selon les auteurs et dans un laps de temps assez réduit : le naïf natif, originel et historique ; le naïf propre à une langue ; le naïf naturel, à valeur esthétique et morale. Chez Sylvius, la notion de naïf s'inscrit dans un mouvement de reconquête des origines, de retour à l'« antique pureté » de la langue française. Le projet est de « décaper » la langue, afin de la rendre pure et simple. Le projet de Louis Meigret est différent : il ne s'agit pas de retourner aux langues anciennes, mais de ramener la variété des usages présents à des règles sous-jacentes. Tout d'abord dans le domaine

de la prononciation et de l'écriture où il cherche à définir la *naïve prononciation* et la *naïve écriture* (qui reconduit sans ambiguïté à sa source orale). De manière distinctive, l'intention de Ramus, Mathieu, ou Henri Estienne dans la seconde moitié du xvi<sup>e</sup> siècle, est la recherche de la singularité du français. Pour H. Estienne, le natif n'a ni lieu ni milieu, c'est davantage une « construction permanente ». Ramus, quant à lui, reconnaît les anomalies et idiotismes comme des « ornements » de la langue. La perspective est différente au xvii<sup>e</sup> siècle puisque le grammairien s'efforce de travailler à fixer la langue française. Le « naïf françois » de Jean Nicot apparaît comme une entreprise de restauration de la pureté linguistique du français, contre les origines étrangères, régionales, ou latines. Charles Maupas, quant à lui, a une vision synchronique de la langue : il écrit une grammaire avec l'intention de rendre compte de la spécificité d'un système linguistique par différenciation avec d'autres, cette grammaire étant à la fois destinée aux apprenants étrangers mais aussi aux amateurs de la langue française. La conception d'une « naïve propriété de notre langue » chez Maupas, loin d'une vision passéiste et conservatrice, cherche à fixer une langue particulière qui est en même temps une langue commune. La définition du naïf par opposition à ce qui l'en éloigne réapparaît chez Antoine Oudin, qui stigmatise l'ancien, mais aussi distingue les « vrais François », qui parlent bien, des locuteurs fautifs. Chez Laurent Chifflet, le naïf définit un usage de la langue en tant que fidèle à la réalité et ne marquant pas d'affectation. Le naïf a basculé, d'une notion appliquée à la langue, elle est devenue une notion appliquée au style.

Maxime Maleux étudie l'argumentation en faveur de l'enseignement de l'hébreu dans le cadre du Collège trilingue de Louvain, institution fondée avec l'aide d'Érasme en 1517 dans le contexte d'un humanisme cherchant à étudier les sources des textes sacrés. Certains théologiens voyaient dans cet enseignement trilingue, comprenant le latin, le grec et l'hébreu, une remise en question de leur autorité sur l'exégèse biblique et attaquent donc sa légitimité. Un des arguments en faveur de l'enseignement de l'hébreu est la *simplicité* de cette langue, argument qui paraît à première vue surprenant. Cette simplicité va avec les idées de pureté, de brièveté et de transparence, et simplement de perfection,

car l'hébreu est alors conçu comme la première langue créée par Dieu. Certains humanistes cherchent à retrouver dans les racines de l'hébreu l'origine des mots dans d'autres langues, car la matérialité de l'hébreu reflète la réalité (de manière cratyliste), et une démarche cabballistique n'est pas possible sans une connaissance de l'hébreu. Quoique cette conception de l'hébreu comme langue première (et universelle) de l'humanité soit remise en doute au xvii<sup>e</sup> siècle, elle a continué à perdurer jusqu'à la fin du xix<sup>e</sup> siècle.

Mat Pires propose un article audacieux et original sur la façon dont la simplicité linguistique a été imaginée dans la science-fiction se déroulant à l'époque préhistorique, dans une gamme de romans, de nouvelles et de bandes dessinées. Un petit nombre de domaines syntaxiques, lexicaux et pragmatiques sont régulièrement ciblés par les auteurs comme lieux de représentation du langage simplifié dans les dialogues de leurs personnages, où Pires voit une homogénéité implicite avec la simplicité imaginée de la vie préhistorique. Ces dialogues sont, bien sûr, présentés non pas dans la langue préhistorique originale, mais dans l'anglais ou le français dans lequel la science-fiction a été écrite, et Pires oppose cette « traduction » au « transfert » des noms de personnes et de peuples, qui sont généralement dans la langue préhistorique imaginée, par exemple *Dorv*, *Zoug*, *Iza*, *Oga* (de Jean M. Auel, *Le clan de l'ours des cavernes*, 1980), bien que parfois « traduits », le « clan de l'ours des cavernes » en étant un tel cas.

Le second fil directeur, sur lequel nous avons organisé les textes de ce volume, est celui de *la comparaison des langues*.

Raf Van Rooy étudie la manière dont la notion de *copia* (« abondance »), empruntée à la rhétorique, organise la réflexion comparative chez les grammairiens du xvi<sup>e</sup> siècle, en servant de critère d'appréciation. Il met d'abord en perspective son étude dans le long terme de l'histoire de l'évaluation de la complexité des langues, entre le consensus sur l'égalité des langues établi au xx<sup>e</sup> siècle et le retour actuelle à la mesurabilité de la complexité des langues, notamment en termes de difficulté et de coût. À la Renaissance, la « copiosité » linguistique est une valeur très positive, et ainsi le grec se retrouve souvent loué par rapport au latin, mais pas chez Scaliger qui, par exemple, trouve l'article du grec superflu et qualifie

les Grecs de peuple excessivement bavard. Van Rooy mène l'enquête sur l'appréciation de *copia* du grec, en examinant les commentaires des grammairiens sur quelques spécificités du grec : l'article, le duel, l'aoriste, les formes participiales, la voix moyenne et son absence d'ablatif (conçue par le grammairien français Cauchie, en faveur du grec, comme une redondance du latin plutôt que comme une richesse !). Enfin, il interroge la superposition des notions de *copia* et de complexité linguistique, et argumente dans le sens d'une filiation possible des deux notions.

Julien Sibileau étudie différents traités de la tradition grammaticale arabe dans lesquels la thématique de la supériorité de l'arabe est développée, en tâchant de séparer entre les arguments théologiques et les arguments de nature plus linguistique. En effet, cette appréciation en faveur de l'arabe apparaît, dès le VIII<sup>e</sup> siècle, comme liée au caractère sacré de la langue : le Coran vante la clarté de la langue arabe, pour le philosophe Abū Ḥātim al Rāzī, il est réputé intraduisible. Dans le traité du X<sup>e</sup> siècle *al-Ṣāhibī fī Fiqh al-Luġa* d'Ibn Fāris, la langue (*kalām*) est approchée en termes de clarté (*bayān*), de système (*luġa*), de compréhension (*fahm*). Ibn Fāris illustre la supériorité de l'arabe en montrant sa richesse lexicale et ses capacités rhétoriques. Chez Ibn Ğinnī, dans un contexte socio-politique qui faisait déjà de la langue arabe la langue de prestige, la vénération de l'arabe par ses locuteurs s'accompagne d'une volonté de la préserver des influences étrangères (notamment du persan). D'autre part la connaissance de l'arabe, selon lui, peut aider au développement de la langue maternelle chez les non-natifs. Ibn Ğinnī reconnaît des traits communs aux langues (*ma'ānī*) qui permettent de comparer les langues, et de les niveler.

Alice Vittrant ré-ouvre le dossier de la simplicité supposée des langues d'Asie. Comme on l'a dit plus haut, on reconnaît généralement la discussion entre Humboldt et Abel-Rémusat comme l'épisode ayant inauguré cette discussion : Humboldt voyant dans les langues flexionnelles un modèle de complexité linguistique et un aboutissement de l'évolution linguistique, Abel-Rémusat lui répondant en développant son argumentation à partir de l'étude du chinois. L'auteure nous montre, à partir d'exemples pris à différentes langues d'Asie du Sud-Est, notamment au birman qu'elle connaît

particulièrement bien, que la complexité des langues d'Asie est un non-vu de certaines théories linguistiques se fondant sur ce qui est « exprimé », car ces langues ont un fonctionnement largement pragmatique, le contexte tenant une place fondamentale dans l'interprétation. De plus, ces langues n'expriment pas forcément les indices personnels, marquent très peu les temps, ce qui, on le comprend, a pu servir d'argument pour conclure à leur simplicité. Vittrant parle en conséquent d'une complexité « cachée », rappelant que la « découverte » des langues asiatiques au XIX<sup>e</sup> siècle avait déjà participé à mettre à mal la focalisation excessive des linguistes sur la morphologie comme révélatrice de leur complexité linguistique. Elle rappelle que la discussion de la mesurabilité de la complexité des langues, à laquelle elle se propose ainsi de participer, à la lumière d'exemples pris dans différentes langues d'Asie du Sud-Est, est un sujet de nouveau admis scientifiquement depuis la fin du XX<sup>e</sup> siècle, après son bannissement dans la période précédente, qui cherchait à défendre l'équité des peuples, des cultures et des langues contre les débordements racistes. On pourrait dire qu'Alice Vittrant, dans une certaine mesure, en répondant aux tenants actuels de la mesurabilité de la complexité linguistique, réitère le geste d'Abel-Résmusat répondant à Humboldt.

Jacques François revient sur un autre épisode, moins connu, de la hiérarchisation des langues : celui de l'approche des langues mandingues par Heymann Steinthal aux prises avec les idées de son époque. François rappelle d'abord que le projet global de Steinthal, exprimé dans le titre de son mémoire d'habilitation, était une synthèse de la linguistique de Humboldt et de la philosophie de Hegel, retenant notamment du premier les idées de « génie de la langue » (*Sprachgeist*) et d'« esprit des peuples » (*Volksggeist*) et du second la dialectique de l'évolution d'âge en âge de l'esprit du monde (*Weltgeist*). Cette vision évolutive va marquer sa typologie linguistique, mais aussi sa vision des langues et des peuples. Son essai sur les langues mandingues (qui a remporté le prix Volney en 1851) se présente d'ailleurs comme une illustration de sa hiérarchisation des langues du point de vue de leur morphologie, celle des langues en question étant défectueuse par rapport aux langues flexionnelles qui représentent l'aboutissement de l'évolution linguistique. François

propose, en fournissant un tableau très utile, une comparaison du classement des types linguistiques proposé par Steinthal en 1860 avec celui de Humboldt (dont Steinthal s'est évidemment inspiré), montrant à la fois les acquis scientifiques (le chinois remonte ainsi dans l'échelle d'évaluation de Steinthal) mais en même temps le point de vue hiérarchisant dont il procède, qui lie systématiquement forme / informe des langues et des représentations qui en découlent. Ceci aboutissant à un bilan plutôt négatif à propos de Steinthal, qu'on imagine habituellement plutôt du côté de Humboldt que de Gobineau, puisqu'il établit par exemple que certains peuples / certaines langues n'ont pas été dotés de la même manière, et qu'il ne reste aux peuples civilisés que le devoir moral de prendre soin des peuples défavorisés... S'ajoute également comme motif de l'évaluation hiérarchisante des langues chez Steinthal, la transposition consciente d'un modèle de représentation de l'évolution biologique et de sa conception de la forme, au domaine linguistique. François propose un tableau reprenant le raisonnement de Steinthal afin de figurer cette transposition du modèle de la morphologie biologique à la morphologie linguistique. Dans l'analyse des langues mandingue, elle apparaît par exemple lorsqu'il conclut à un « manque d'isolation des représentations » dans la langue soso, ceci renvoyant à l'image de la monade, un élément biologique indépendant et isolable.

José Edicarlos De Aquino analyse les motifs de la hiérarchisation des langues dans les *Traços Geraes de Linguistica* (1880) de Júlio Ribeiro, ouvrage considéré comme le premier ouvrage de linguistique au Brésil, ouvrage dont l'approche est largement marquée par un questionnement sur l'origine du langage et sur son développement, et dont l'un des objectifs est de proposer une typologie des langues. De Aquino situe le projet de Ribeiro par rapport aux théories sur l'origine du langage au XIX<sup>e</sup> siècle et en particulier par rapport aux recherches françaises mêlant linguistique et anthropologie, en particulier Henri Lefèvre qui lui sert de modèle pour penser les différents stades de l'évolution du langage et de l'humanité. Ribeiro imagine sept périodes successives, les premières préhistoriques (traduction des impressions en cris de douleur ou de joie, usage démonstratif en vue de la communication, invention des mots par généralisation des impressions), puis des périodes

historiques (transformation des éléments antérieurs en substantifs et en verbes ; puis se succèdent les périodes monosyllabique, agglutinative amalgamante et contractée). Ces stades du développement des langues correspondent à sa typologie linguistique : ainsi le chinois est rangé dans la catégorie des langues monosyllabiques, l'étape la plus ancienne et la moins élaborée du développement des langues modernes. Des explications politiques, sociales, économiques (dont la « division du travail »), artistiques, scientifiques sont également données pour expliquer l'évolution favorable des langues.

Dans la troisième partie, il s'agit de *la langue et la pensée*. Jusqu'ici nous avons traité du deuxième millénaire, avec parfois un plongeon dans le troisième. Le chapitre de Maria Chriti repousse cependant la période de sept siècles, montrant comment les écrits d'un philosophe néoplatonicien du troisième siècle, Porphyre de Tyr, construisent un récit de l'origine et de l'évolution du langage qui repose sur un progrès évolutif à partir d'une simplicité initiale, aboutissant à la complexité, avec l'implication d'un progrès intellectuel correspondant. Cette correspondance présumée finira par s'avérer être le talon d'Achille du modèle, en particulier à la fin du xx<sup>e</sup> siècle et après, lorsque les implications sur la simplicité seront largement exposées comme fondées sur des hypothèses préjudiciables. L'œuvre de Porphyre a servi de base à l'école néoplatonicienne d'Alexandrie aux v<sup>e</sup> et vi<sup>e</sup> siècles, et cette continuité historique suggère fortement qu'il faut accorder encore plus d'attention à ce penseur qui, bien qu'il ne soit certainement pas oublié, tend à être si étroitement associé à des idées et positions particulières, que ses idées vis-à-vis de la simplicité linguistique ont été marginalisées.

Bien qu'à la fin du xx<sup>e</sup> siècle le lien langage-esprit se dissocie des comparaisons intellectuelles, il ne disparaît pas – au contraire, il assume l'une de ses manifestations historiques les plus fortes dans la grammaire transformationnelle-générative de Noam Chomsky. L'article de Jean-Michel Fortis étudie la relation complexe qui existait à l'époque chomskyenne entre la linguistique et la psychologie, plus particulièrement de la « psycholinguistique transformationnelle » qu'il caractérise comme entrant dans un bref mariage avec la grammaire générative, suivi d'un « divorce à l'américaine ». Fortis

construit un cadre dans lequel un large éventail de théories et de méthodologies syntactiques, dont certaines ne sont généralement pas classées comme génératives ou psycholinguistiques, peuvent être analysées comme des réponses aux tensions créées par ce conflit de sous-champs, y compris les différences doctrinales concernant la complexité des transformations et du « processing ».

*La complexité et la société* est le sujet des deux chapitres suivants, qui portent sur la manière dont les conceptions linguistiques de la simplicité et de la complexité ont fait l'objet de tentatives d'applications pédagogiques dans un contexte sociologique. En Grande-Bretagne, Basil Bernstein est l'auteur ayant le plus cherché à articuler la complexité de la langue et ses conséquences pédagogiques. Ses travaux sur les codes « *restricted* » et « *elaborated* » faisaient partie d'un programme pour améliorer l'éducation des enfants de la classe ouvrière, en sensibilisant les enseignants au fait que la langue familiale de ces élèves manquait, selon Bernstein, de la complexité de leurs camarades de couche moyenne et supérieure. Marie-Madeleine Bertucci fait ressortir le contraste entre l'approche de Bernstein et celle de William Labov, qui en a pris le contre-pied en plaidant pour l'acceptation du Black English, l'idiome de la plupart des Noirs américains, comme une langue à part entière, plutôt qu'une déformation de l'anglais standard. Ici la cible est encore une fois la sensibilisation des enseignants et des institutions dans lesquelles ils travaillaient. Labov a démontré que le Black English est gouverné par des règles, qui ne sont pas moins complexes que celles de la langue standard, voire parfois plus. Il a grandement contribué à l'acceptation du Black English dans le système éducatif américain.

Bernstein figure également en bonne place dans le chapitre de John E. Joseph, aux côtés de son camarade linguiste Michael Halliday, qui a beaucoup fait pour promouvoir les conceptions de Bernstein en linguistique appliquée. Joseph compare et oppose leur approche au type de simplification structurelle que l'on trouve chez les syntacticiens Lucien Tesnière et Noam Chomsky. Le chapitre contient une correspondance inédite entre Tesnière et un instituteur français qui a appris à ses élèves à effectuer l'analyse tesnièreenne par les « stemmas », dans laquelle même les phrases les plus compliquées ont un centre « solaire » auquel tout se connecte. Le générativisme

transformationnel de Chomsky était motivé par une simplification dans le même esprit, réduisant les structures de phrases et les classes de mots jusqu'au plus petit dénominateur commun, et il a suscité de nombreuses tentatives d'application par des pédagogues qui croyaient faire avancer la vision chomskyenne, inspirée par Humboldt, de la « créativité linguistique infinie » des enfants.

Les deux derniers chapitres concernent des tentatives de *mesure quantitative de la complexité linguistique*. Margaret Thomas poursuit le thème pédagogique, en proposant la première étude historique du développement de la « T-unit », une mesure de la « maturité » syntaxique – autrement dit de la complexité – créée par le pédagogue américain Kellogg W. Hunt. Thomas montre les liens qu'avait la T-unit avec la linguistique théorique des années 1950 et 1960, pour laquelle il a offert un pont aux pédagogues cherchant à mettre en pratique certains aspects de la syntaxe générative, par exemple. La T-unit n'est pas entièrement oubliée en linguistique appliquée et en psychologie de l'éducation, mais pour la plupart des lecteurs de l'article de Thomas, ce sera une révélation.

Enfin, Vassilis Symeonidis et Nikolaos Lavidas passent en revue les tentatives de mesure de la complexité du travail des linguistes historiques des dernières décennies, ainsi que les controverses que ces tentatives ont soulevées. En se concentrant sur le processus de grammaticalisation, les auteurs proposent leur propre dichotomie entre complexité « structurelle » et complexité « dynamique », où la première représente le mode plus traditionnel d'énumérer les structures et les options que la grammaire contient ou autorise, et la seconde une démarche diachronique montrant la façon dont le « retranchement » progressif des structures nouvellement créées rétablit l'équilibre de complexité que l'innovation initiale a perturbé.

## Bibliographie

- Joseph, John E. & Frederick J. Newmeyer. 2012. "All languages are equally complex": the rise and fall of a consensus. *Historiographia Linguistica* 39(2-3). 341-368.

Lovejoy, Arthur O. & George Boas. 1935. *Primitivism and related ideas in Antiquity*. With supplementary essays by W. F. Albright and P. E. Dumont. Baltimore : Johns Hopkins Press; London: Oxford University Press, Humphrey Milford.

Le naturel, le simple et le natif



# L'évolution du concept d'ordre « naturel » à partir du XVII<sup>e</sup> siècle

**Gerda Hassler**

Université de Potsdam

---

## RÉSUMÉ

Le développement de la notion d'un ordre fixe des mots qui était conçue comme un critère positif dans l'évaluation des qualités d'une langue aux XVII<sup>e</sup> et au XVIII<sup>e</sup> siècles est bien décrit pour le français. La présente contribution examinera cette évolution du point de vue de la simplification de la langue. La discussion sur l'ordre naturel ne pouvait pas manquer de se poursuivre dans d'autres pays et durant les siècles suivants. On peut observer des continuités linéaires de la doctrine de l'*ordre naturel* qui menaient même à exiger d'adapter l'ordre des mots d'autres langues (entre autres, de l'anglais, de l'allemand, de l'espagnol, du russe) à la construction directe. Dans des théories contemporaines, fondées sur le rationalisme, on retrouve la même association de la simplicité et de l'ordre naturel.

## MOTS-CLÉS

ordre naturel, simplification de la langue, normalisation, inversion, rationalisme

## ABSTRACT

The development of a notion of fixed word order that was conceived as a positive criterion in the evaluation of languages in the 17th and 18th centuries is well described for French. This contribution will examine this development from the point of view of language simplification. The discussion on natural order could not fail to continue in other countries and in the centuries that followed. Linear continuations of the doctrine of *natural order* can be observed, which even led to a demand to adapt the word order of other languages (among others, English, German, Spanish, Russian) to direct con-

struction. In modern theories based on rationalism, the same combination of simplicity and natural order can be found.

#### KEYWORDS

natural order, language simplification, normalisation, inversion, rationalism

---

## 1. Introduction

L'hypothèse selon laquelle un certain ordre de mots est favorable à la communication et au traitement de l'information est commune à toutes les théories qui s'occupent de ce sujet, à partir de la postulation d'un ordre fixe qui était déjà conçue comme un critère positif dans l'évaluation des qualités d'une langue par Louis Meigret (ca. 1500-ca. 1558) jusqu'aux théories modernes de la structure informationnelle. D'un côté, certains théoriciens considèrent que la simplicité et le fait de coûter moins d'effort dans le traitement de l'information coïncident avec un ordre fixe consistant en l'arrangement du sujet avant le verbe et de l'objet après celui-ci ; de l'autre, des contradicteurs affirment que l'arrangement des mots doit suivre l'intention communicative du locuteur pour être décodé avec moins d'effort. Ainsi, dans un article écrit dans le cadre du programme minimaliste, on peut trouver l'alinéa suivant :

*Much of the current literature on scrambling is based on the assumption that fixed constituent order is something one gets for free, while free constituent order requires extra efforts, viz. reordering transformations. The purpose of this section is to show that this conviction is wrong, and it will sketch three descriptions of free word order that are compatible with minimalism.* (Fanselow 2003 : 191)

'Une grande partie de la littérature actuelle sur le scrambling [litt. brouillage] est basée sur l'hypothèse que l'ordre des constituants fixe est quelque chose que l'on obtient gratuitement, alors que l'ordre des constituants libre exige des efforts supplémentaires, à savoir le réarrangement des transformations. Le but de cette section est de montrer que cette conviction est fautive, et elle esquisse trois descriptions de l'ordre constitutif libre qui sont compatibles avec le minimalisme.'

L'ordre des mots qui coûterait le moindre effort de traitement est souvent appelé *ordre naturel*. Mais que signifie le fait qu'un ordre

soit déclaré comme naturel et plus simple pour le traitement de la langue ? Il n'est pas possible de donner une réponse qui soit valable pour tous les auteurs et à toutes les époques, le contenu de cette notion ayant changé en fonction des positions philosophiques, grammaticales ou rhétoriques. Nous analyserons d'abord l'usage du concept d'*ordre naturel* et traiterons sa signification dans différents courants et contre-courants grammatico-philosophiques. Ensuite, une attention particulière sera accordée aux tentatives de normalisation de l'ordre des mots visant à rendre la langue plus facile à utiliser et à minimiser les coûts de traitement. En dernier lieu, nous examinerons la réapparition du concept de l'ordre naturel dans la discussion moderne de la structure informationnelle.

## 2. Le concept de l'ordre naturel et la simplicité de la langue

Une théorie d'un ordre de mots provenant de la nature peut déjà être trouvée dans les écrits de Denys d'Halicarnasse (vers 60 av. J.-C. – 8 av. J.C.). Cependant, ce dernier ne préconise pas un ordre de mots particulier, mais justifie plutôt divers ordres de mots. Quintilien (35-96) mentionne l'hypothèse d'un ordre naturel des mots en utilisant les désignations *rectus ordo*, *naturalis ordo*, pour ensuite les rejeter. Sous *naturalis ordo*, il comprend une séquence fixe dans des paires de mots comme *le jour et la nuit*, *la terre et la mer*, mais il reconnaît aussi des déviations de celle-ci pour des raisons rythmiques (*transgressio* ou *hyperbaton*). Cicéron (106 av. J.-C. – 43 av. J.-C.) utilise également ces désignations pour traiter les fonctions rhétoriques de l'ordre des mots. Pour les rhétoriciens du IV<sup>e</sup> siècle, l'arrangement des pensées en disposition libre est caractéristique et l'*ordo naturalis* est opposé à l'*ordo artificialis*.

Comme beaucoup de concepts, le concept d'ordre naturel des mots est transféré de la rhétorique à la grammaire. L'élaboration d'un concept d'ordre de mots nécessaire et fixe qui est évalué comme un critère positif de la qualité de la langue est généralement attribuée en France à Meigret qui, dans son *Treuvé de la grammaire françoise* (1550, 142-143), déclare comme nécessité générale le maintien d'un ordre de mots fixe, qu'il appelle *la construccïon ou ordonance*

*bone de parolles*. Attribuer une valeur exemplaire au latin avec ses nombreuses inversions signifierait ne pas reconnaître l'originalité et la valeur propre du français.

Une nouvelle étape dans le développement de la doctrine de l'ordre naturel est initiée par les grammairiens de Port-Royal, Antoine Arnauld (1612-1694) et Claude Lancelot (1615/16-1695) (cf. entre autres Chevalier 2006 [1968], Delesalle 1980, Ricken 1978, 1994, Seguin 1993, Haßler 2012, 2014). Dans ses manuels de langue, Lancelot distingue une *construction simple ou régulière*, qui correspond le plus à l'usage de la langue vernaculaire, d'une construction *irrégulière ou figurée*, qui s'en écarte très légitimement dans un but précis de brièveté ou d'élégance (cf. Lancelot 1650 [1644] et 1656 [1655]). La grammaire de Port-Royal (1660) permet également de s'écarter de l'ordre fixe des mots à des fins communicatives. Le programme de la grammaire générale, qui vise à étudier les lois générales de la pensée et postule la raison comme principe explicatif d'un usage apparemment inexplicable du langage, suggère une interprétation de l'ordre naturel au sens du rationalisme de René Descartes (1596-1650) :

*Des figures de construction.* Ce que nous avons dit ci-dessus de la syntaxe, suffit pour en comprendre l'ordre naturel, lorsque toutes les parties du discours sont simplement exprimées, qu'il n'y a aucun mot de trop ni de trop peu, & qu'il est conforme à l'expression naturelle de nos pensées.

Mais parce que les hommes suivent souvent plus le sens de leurs pensées, que les mots dont ils se servent pour les exprimer, & que souvent, pour abréger, ils retranchent quelque chose du discours, ou bien que regardant à la grace, ils y laissent quelque mot qui semble superflu, ou qu'ils en renversent l'ordre naturel ; de-là est venu qu'ils ont introduit quatre façons de parler, qu'on nomme *figurées*, & qui sont comme autant d'irrégularités dans la Grammaire, quoiqu'elles soient quelquefois des perfections & des beautés dans la Langue. (Arnauld & Lancelot 1768 [1660] : 234-235)

D'autres auteurs, tels que Géraud de Cordemoy (1626-1684) (1668) et Louis Le Laboureur (1615-1679) (1669) contribuent à une justification de l'ordre naturel sur une base rationaliste. La conception de l'ordre naturel sur laquelle se fondent ces rationalistes consiste à supposer que l'ordre établi par la langue française dans la séquence sujet-verbe-objet suit l'ordre dans la pensée et permet donc de la reconstruire plus facilement. Considérant la séquence du

développement des éléments individuels de la langue, Cordemoy en vient à la conclusion que leur appropriation dans l'acquisition de la langue et leur disposition en français ont une base commune. Lorsque Le Laboureur pose la question de savoir si les grandes différences entre l'ordre des mots latins et français signifient que les Romains pensent différemment des Français, il répond en arguant de la séparation rationaliste de la langue et de la pensée. Selon lui, la raison, et donc les lois de la pensée, sont communes à tous les peuples. Le Laboureur ne doute aucunement de l'unité et de l'indivisibilité de la raison :

J'ai dit que les Latins pensaient ainsi que nous autres Français [...]. (Le Laboureur 1669 : 149)

Comme leur tête n'était point faite autrement que la nôtre, il y a grande apparence qu'ils concevaient les choses de la même façon que nous les concevons ; et cela est d'autant plus vraisemblable que notre Logique n'est point différente de la leur. (Le Laboureur 1669 : 167)

Nous suivons en tous nos discours exactement l'ordre de la pensée qui est celui de la Nature ; [...] l'usage et la coutume ne sauraient nous imposer en cela, où la raison se fait presque sentir et toucher. (Le Laboureur 1669 : 173)

Le terme de *clarté* désigne d'abord l'organisation judicieuse et compréhensible des idées dans la structure du discours, qui doit avoir pour objectif de rendre l'intention du locuteur clairement reconnaissable et compréhensible pour le public. La notion de *clarté* fut toutefois réinterprétée comme une qualité systématique de certaines langues. Dans le cadre de l'affirmation du rationalisme, cette qualité de clarté est revendiquée avant tout pour le français et décrite comme caractéristique de cette langue. La notion de « clarté » est associée au concept cartésien de la preuve, selon lequel il y a des vérités primaires et évidentes qui n'exigent aucune preuve, mais qui peuvent être utilisées dans l'argumentation. Dans la communication, l'utilisation générale de l'ordre naturel en français facilite la compréhension et le traitement de l'information. En revanche, les écarts par rapport à cette règle rendent la compréhension plus difficile en raison de la reconstruction supplémentaire requise de la séquence naturelle de la pensée. Selon Charpentier, c'est la raison pour laquelle le français a une simplicité naturelle et un pouvoir de

persuasion plus grand que le latin qui, avec son mode d'expression artificiel, soulève des doutes quant à la sincérité :

Chapitre XXX. Que la Construction Directe comme est celle de la Langue Française, est incomparablement plus estimable que la Construction Renversée de la Langue Latine, Et que les Grecs & les Latins mesmes en ont jugé de la sorte. (Charpentier 1683 : 639)

Ainsi on ne doit pas douter que la plus belle Elocution, & mesme la plus heureuse, ne soit celle ou sans renverser l'Arrangement Naturel des paroles, elles se joignent agreablement ensemble, & forment une Cadence pleine d'harmonie. (Charpentier 1683 : 645)

Il n'y a que la Langue Française qui puisse exprimer les grandes émotions, & converser la politesse d'un Discours estudié, parce que dans cette Langue la Nature & l'Art sont presque toujours d'accord. Ses Nombres ne paroissent point recherchez, parce qu'ils ne s'éloignent jamais de la Construction Naturelle, & quoy que cette Construction n'empesche pas qu'on n'en sente l'Harmonie, elle fait neantmoins que l'Art en est plus caché, & par consequent qu'elle est plus propre à persuader. Mais le principal avantage de la Construction Directe, c'est qu'elle contribue à la Clarté, qui est la souveraine perfection du Discours. (Charpentier 1683 : 650)

Ces théories rationalistes attribuent à l'ordre naturel du français la capacité de rendre plus facile le traitement cognitif des énoncés parce qu'il n'y a pas de décalage entre la langue et la pensée. Cette idée a aussi des conséquences pour l'enseignement du latin, langue prétendue à inversions. Pour rendre cet apprentissage plus facile aux élèves, les constructions latines jugées compliquées sont d'abord adaptées à l'ordre naturel des mots. Un exemple des *Véritables principes de la grammaire, ou nouvelle grammaire raisonnée pour apprendre la langue latine* (1729) de Du Marsais, qui est déjà utilisé par Joseph de Jouvancy en 1692 (1643-1719), illustre cette approche :

Minóis filiam Ariádnen, cujus ope labyrínthi ambáges explicúerat, secum ábstulit : eam tamen, immemor beneficii, deséruiť in insulá Naxo : destitúťam Bacchus duxit.

ábstulit secum Ariádnen, filiam Minóis ; ab ope cujus  
*il emmena avec lui Ariane, fille de Minos par le secours de laquelle*  
 et enleva Ariane, fille de Minos. Cète princesse avoit donné à Thésée

explicúerat ambáges labyrinthi. Tamen, immemor  
*il avoit démêlé les détours du labyrinthe. Cependant, ne se ressouvenant point*  
 un peloton de fil qui aida ce héros à sortir du labirinte (\*) Cependant, oubliant  
 (\*) Où il étoit entré pour tuer Minotaure.

beneficii, deséruiet eam in insulá Naxo, Bacchus  
*de ce bienfait, il l'abandona dans l'île Naxe, où Baccus*  
 le service qu'elle lui avoit rendu, il l'abandona dans l'île de Naxe, où Baccus

duxit in domum Ariáden destitútam à Théseo.  
*emmena dans sa demeure Ariane abandonée par Thésée.*  
 l'épousa.

(Du Marsais 1797 [1729], II : 215)

Alors que la doctrine de l'*ordre naturel* se nourrit de la discussion philosophique de l'époque et qu'elle repose sur l'hypothèse d'une pensée innée et incorporelle, on aurait pu constater, par l'observation de la pratique linguistique, qu'un ordre fixe des mots ne correspond pas à la réalité. Déjà Le Laboureur présente la simultanéité des perceptions comme une objection à la théorie de l'ordre naturel (Ricken 1978 : 65). On pourrait dire, lorsqu'on se représente un chien en train d'attraper un lièvre qu'il s'agit là d'un ensemble de représentations simultanées, comme si on voyait un tableau. Il est pourtant certain que le locuteur, s'il veut exprimer l'action par l'activité du chien, doit d'abord penser au chien, seulement ensuite à son activité et à l'objet de celle-ci. Si, par contre, on exprime l'action par la souffrance du lièvre en disant « le lièvre est pris par le chien », on pense d'abord au sujet et ensuite à l'action, de sorte que cette phrase, elle aussi, observe l'ordre naturel (Le Laboureur 1669 : 168-170).

On trouve un exemple d'explication rhétorique et pragmatique de l'ordre des mots chez Bernard Lamy (1640-1715) qui part d'un programme rationaliste et métaphysique, mais qui admet, dans sa *Rhétorique*, que l'*usage* est un principe régulateur qui agit bien souvent contre les principes de la *raison*, comprise comme une qualité innée de l'être humain. La *raison* elle-même exige, cependant, d'observer les principes de l'usage parce que sans ceux-ci la compréhension serait impossible. Ainsi, Lamy exige tout d'abord

qu'on suive la *nature*, qu'on exprime simplement ses idées. Les signes de celles-ci sont *nécessaires* et les plus grandes qualités d'un discours sont sa *netteté* et sa *naïveté*. Par *netteté* et *naïveté*, on entend alors la simplicité et la clarté du discours, qui doit exprimer la pensée sans digression, sans fioritures baroques et sans inversion dans l'ordre des mots. Dans ce sens, Lamy exige qu'on *suive l'ordre le plus naturel, le plus simple* :

Lorsque toutes nos idées sont exprimées avec leur liaison, il est impossible que l'on n'aperçoive ce que nous pensons, puisque nous en donnons tous les signes nécessaires. C'est pourquoi ceux-là parlent clairement qui parlent simplement, qui expriment leurs pensées d'une manière naturelle, dans le même ordre, dans la même étendue qu'elles ont dans leur esprit. [...] Nôtre langue ne se sert point de cette figure, ni de toutes les autres figures de Grammaire, elle aime la netteté & la naïveté ; c'est pourquoi elle exprime les choses autant qu'il se peut dans l'ordre le plus naturel, le plus simple. (Lamy 1688 [1675] : 44-45)

Au milieu du XVIII<sup>e</sup> siècle, plusieurs auteurs sensualistes contribuent à une considération de la linéarisation de la pensée par le langage dans une perspective plus large, qui n'est plus restreinte au débat sur l'ordre naturel. Ils tiennent compte de questions épistémologiques aussi bien que de la pertinence de problèmes esthétiques et – comme on dirait aujourd'hui – pragmatiques. Étienne Bonnot de Condillac (1714-1780) discute les inversions dans l'*Essai sur l'origine des connoissances humaines* (1746), Charles Batteux (1730-1780) les étudie sous un angle esthétique dans ses *Lettres sur la phrase française comparée avec la phrase latine* (1748), et, en 1751, paraît la *Lettre sur les sourds et muets* de Denis Diderot (1713-1784) comme une réponse à Batteux.

Condillac justifie son refus de la doctrine de l'*ordre naturel* historiquement, en le mettant en relation avec sa conception sensualiste du développement du langage et de la pensée. Pour Condillac, *ce qu'on appelle ici naturel varie nécessairement selon le génie des langues* et dépend du degré de leur fixation et normalisation :

*Des inversions.* [N]ous nous flattons que le françois a, sur les langues anciennes, l'avantage d'arranger les mots dans le discours, comme les idées s'arrangent d'elles-mêmes dans l'esprit; parce que nous nous imaginons que l'ordre le plus naturel demande qu'on fasse connoître le sujet dont on parle, avant d'indiquer ce qu'on en affirme ; c'est-à-dire, que le

verbe soit précédé de son nominatif et suivi de son régime. Cependant nous avons vu que, dans l'origine des langues, la construction la plus naturelle exigeoit un ordre tout différent. Ce qu'on appelle ici naturel varie nécessairement selon le génie des langues, et se trouve dans quelques-unes plus étendu que dans d'autres. (Condillac 1961 [1746] : 164)

Ce qu'on appelle *construction naturelle* dépend alors de la situation historique de la langue. Ainsi, quand les hommes ont commencé à utiliser des sons articulés, « le nom de la chose se présenta naturellement le premier, comme étant le signe le plus familier » (Condillac 1961 [1746] :123). Comme le langage d'action servait de modèle au langage articulé, celui-ci « a dû, dans les commencemens, conserver les idées dans le même ordre que l'usage du premier avoit rendu le plus naturel » (Condillac 1961 [1746] :124) :

Cette manière de s'énoncer étoit la plus commode pour celui qui parloit et pour celui qui écoutoit. Elle l'étoit pour le premier, parce qu'elle le faisoit commencer par l'idée la plus facile à communiquer : elle l'étoit encore pour le second, parce qu'en fixant son attention à l'objet dont on vouloit l'entretenir, elle le préparoit à comprendre plus aisément un terme moins usité, et dont la signification ne devoit pas être si sensible. Ainsi l'ordre le plus naturel des idées vouloit qu'on mît le régime avant le verbe : on disoit, par exemple, *fruit vouloir*. (Condillac 1961 [1746] :123-124)

Aussi, au début des langues articulées, doit-on commencer par prononcer le régime qui exprime normalement l'objet voulu :

Le verbe venant après son régime, le nom qui le régissoit, c'est-à-dire le nominatif, ne pouvoit être placé entre deux ; car il en auroit obscurci le rapport. Il ne pouvoit pas non plus commencer la phrase, parce que son rapport avec son régime eut été moins sensible. Sa place étoit donc après le verbe. Par-là, les mots se construisoient dans le même ordre dans lequel ils se régissoient ; unique moyen d'en faciliter l'intelligence. On disoit *fruit vouloir Pierre*, pour *Pierre veut du fruit* ; et la première construction n'étoit pas moins naturelle que l'autre l'est actuellement. (Condillac 1961 [1746] : 124)

La construction française *Alexandre a vaincu Darius* n'est pas plus naturelle que la construction latine *Darium vicit Alexander*, c'est le manque de déclinaisons morphologiques qui ne permet pas de concilier la clarté avec un ordre différent :

Le latin en est la preuve ; il allie des constructions tout-à-fait contraires, et qui néanmoins paroissent également conformes à l'arrangement des idées. Telles sont celles-ci : *Alexander vicit Darium, Darium vicit Alexander*. Si nous n'adoptons que la première, *Alexandre a vaincu Darius*, ce n'est pas qu'elle soit seule naturelle ; mais c'est que nos déclinaisons ne permettent pas de concilier la clarté avec un ordre différent. Sur quoi seroit fondée l'opinion de ceux qui prétendent que dans cette proposition, *Alexandre a vaincu Darius*, la construction française seroit seule naturelle ? Qu'ils considèrent la chose du côté des opérations de l'ame, ou du côté des idées, ils reconnoîtront qu'ils sont dans un préjugé. En la prenant du côté des opérations de l'ame, on peut supposer que les trois idées qui forment cette proposition se réveillent, tout à la fois, dans l'esprit de celui qui parle, ou qu'elles s'y réveillent successivement. Dans le premier cas, il n'y a point d'ordre entr'elles : dans le second, il peut varier, parce qu'il est tout aussi naturel que les idées d'Alexandre et de vaincre se retracent à l'occasion de celle de Darius ; comme il est naturel que celle de Darius se retrace à l'occasion des deux autres. L'erreur ne sera pas moins sensible, quand on envisagera la chose du côté des idées : car la subordination qui est entr'elles autorise également les deux constructions latines ; *Alexander vicit Darium, Darium vicit Alexander* : en voici la preuve. Les idées se modifient dans le discours, selon que l'une explique l'autre, l'étend, ou y met quelque restriction. Par-là, elles sont naturellement subordonnées entr'elles ; mais plus ou moins immédiatement, à proportion que leur liaison est elle-même plus ou moins immédiate. (Condillac 1961 [1746] : 164-166)

De plus, Condillac propose, avec la notion de la *liaison des idées*, un concept synchronique qui doit permettre le fonctionnement de la langue. En lieu et place de l'*ordre naturel* rationaliste, il pose la *liaison des idées* comme principe qui lie toutes les parties du discours.

Charles Batteux (1713-1780) établit un nouveau dogme, opposé à celui de l'ancienne doctrine de l'ordre naturel, en affirmant qu'il n'y a pas d'ordre naturel des mots français mais, au contraire, des inversions partout :

Inversement, en général ; signifie renversement. Ainsi, quand on demande, si dans une langue il y a des inversions, c'est demander, s'il s'y fait des renversemens. Mais quels sont-ils ces renversemens qu'on peut faire ? N'y en a-t-il que d'une sorte ? Si tout renversement suppose un ordre contraire à celui qui est renversé, & qu'on puisse distinguer plusieurs de ces ordres dans ce qu'on appelle langue ; il doit y avoir aussi plusieurs espèces de renversemens. Or il y a 1., l'ordre des pen-

sées : 2., l'ordre des expressions ; 3., l'ordre d'une langue particulière, soit par opposition à quelque autre langue, avec laquelle on s'avise de la comparer ; soit en la comparant avec elle-même, dans les deux genres de langage qu'elle contient, qui sont la Prose & la Poésie. (Batteux 1763 : 8)

Lorsqu'on veut mettre en évidence les qualités particulières d'un objet, la place de l'adjectif est, selon Batteux, avant le substantif. Ainsi dira-t-on en présence d'un Romain *Romanum imperium* et non *imperium Romanum* (Batteux 1747-1748, II : 18). Batteux distingue de cet ordre *naturel pratique* un autre ordre qu'il appelle *spéculatif* ou *métaphysique* et qui correspond à une manière de penser tout à fait abstraite. En proclamant l'ordre pratique comme ordre naturel, Batteux inverse la théorie rationaliste et établit un nouveau dogme sensualiste.

Diderot, en revanche, propose l'étude systématique de la langue des gestes dont l'arrangement parallèle des pensées devrait correspondre à celui du premier stade de l'humanité avant la formation du langage articulé. Les hypothèses de Diderot permettent de conclure que le français présente beaucoup de changements par rapport à l'*ordre naturel*. Le terme *ordre naturel* ne désigne pas, dans ce cas, l'ordre sujet-verbe-objet, mais l'ordre de la nature, c'est-à-dire de l'origine du langage.

Le désir d'avoir des normes pour l'arrangement des mots mène à la formation de normes syntaxiques. Mettant en doute leur caractère naturel, Diderot ne les appelle pas *ordre naturel*, mais *ordre d'institution*, *ordre scientifique*, *ordre grammatical* ou *ordre didactique* :

Je dis l'*ordre naturel* des idées ; car il faut distinguer ici l'*ordre naturel* d'avec l'*ordre d'institution*, et, pour ainsi dire, l'*ordre scientifique* : celui des vues de l'esprit, lorsque la langue fut tout à fait formée. (Diderot 1969 [1751] : 523)

Un élément remarquable dans la théorie de Diderot est la définition nettement pragmatique de toutes les notions, y compris celle de l'ordre naturel. Il dit ainsi que, dans une phrase isolée qui ne renferme qu'un petit nombre d'idées, « il est fort difficile de déterminer quel est l'ordre naturel que ces idées doivent avoir par rapport à celui qui parle » (Diderot 1969 [1751] : 539). Il se pose même la question de savoir si l'esprit ne peut pas en avoir un certain nombre au même moment. Cela montre qu'il considère la linéarisation comme une

contrainte qu'il vaudrait mieux éviter pour garder la simultanéité des pensées :

Mais allons plus loin : je soutiens que, quand une phrase ne renferme qu'un très petit nombre d'idées, il est fort difficile de déterminer quel est l'ordre naturel que ces idées doivent avoir par rapport à celui qui parle ; car si elles ne se présentent pas toutes à la fois, leur succession est au moins si rapide, qu'il est souvent impossible de démêler celle qui nous frappe la première. Qui sait même si l'esprit ne peut pas en avoir un certain nombre exactement dans le même instant ? (Diderot 1969 [1751] : 539)

En poursuivant dans une perspective historique, Diderot suppose que les catégories d'Aristote ont eu une influence sur la fixation de l'ordre des mots en français. À l'époque où l'ordre des mots – beaucoup plus libre auparavant – se normalise, la doctrine d'Aristote pèse d'un grand poids dans le système scolaire en France.

Le contenu du concept d'« ordre naturel » change donc aux XVII<sup>e</sup> et XVIII<sup>e</sup> siècles. Alors que chez les auteurs rationalistes du XVII<sup>e</sup> siècle, il signifie l'assimilation à l'ordre présupposé dans la pensée, les penseurs sensualistes y opposent leur conception de la profération selon l'ordre d'importance pour le locuteur. Pour Diderot en particulier, il n'est pas possible de déterminer le caractère naturel de l'ordre des mots dans une phrase isolée, car il dépend de l'intention respective de chaque locuteur.

### 3. L'ordre naturel dans la normalisation des langues

L'ordre supposé « naturel » sujet-*verbe*-*objet* du français qui n'est pas l'ordre des origines, est attribué par des grammairiens-philosophes à la normalisation de la langue et à son abstraction. Examinons maintenant les effets du concept d'« ordre naturel » dans la normalisation des langues.

L'ordre fixe des mots français est un élément déjà couramment utilisé dans les efforts de normalisation du XVII<sup>e</sup> siècle. En 1618, le grammairien Charles Maupas (1570-1625) fait remarquer que dans son ordre des mots, la langue française suit les règles de la logique :

Nostre langue aime à suivre, en l'arrangement de nos mots, l'ordre naturel de l'entendement, qui est la diction regissant soit devant la regie. (Maupas 1632 [1618] : 251-252)

Claude Favre de Vaugelas (1585-1650) va même plus loin, en déclarant que le respect de l'ordre naturel est une condition préalable à la clarté du langage. Même Dominique Bouhours (1628-1702) qui, en tant que disciple de Vaugelas, déclare l'indépendance de l'usage de la langue par rapport à la raison, se trouve contraint de faire une exception pour l'ordre naturel des mots et devient partisan du modèle d'explication rationaliste. Selon lui, ce qui est décisif pour relier les mots, c'est la raison : elle établit une séquence naturelle qui n'est suivie que par la langue française. Alors que les autres langues doivent violer l'ordre naturel pour atteindre l'harmonie et l'élégance, le français n'a qu'à suivre l'ordre naturel pour satisfaire aux exigences esthétiques :

C'est que la langue françoise est peut-estre la seule, qui suive exactement l'ordre naturel, & qui exprime les pensées en la maniere qu'elles naissent dans l'esprit. *Je m'explique, & je vous prie de m'entendre.* Les Grecs & les Latins ont vn tour fort irregulier ; pour trouver le nombre & la cadence, qu'ils cherchent avec tant de soin, ils renversent l'ordre dans lequel nous imaginons les choses : ils finissent le plus souvent leurs periodes, par où la raison veut qu'on les commence.

Le nominatif qui doit estre à la teste du discours selon la regle du bon sens, se trouve presque toujours au milieu ou à la fin. Par exemple, au lieu de dire naturellement & régulièrement comme nous, *César a vaincu Pompée dans la bataille de Pharsale* : ils disent en leur langage, *de Pharsale dans la bataille a vaincu Pompée César.* (Bouhours 1671 : 57-58)

Dans d'autres traditions nationales de normalisation des langues, où le point de départ est différent, l'influence de la doctrine rationaliste se fait aussi sentir. Il n'est guère étonnant que le problème de l'ordre des mots soit évoqué dans la grammaire d'Antonio de Nebrija (1441/44-1522) : cet ordre est complètement différent en espagnol et en latin. Nebrija traite de deux aspects du problème : (1) la relation que certains éléments établissent entre eux, par exemple l'adjectif et le substantif ou le substantif et le verbe, et (2) la relation de l'ordre des mots avec les conditions extra-linguistiques :

*Este concierto de las partes de la oración entre sí es natural a todas las naciones que hablan, porque todos conciertan el adjetivo con el substan-*

*tivo, y el nominativo con el verbo, y el relativo con el antecedente ; mas así como aquestos preceptos son a todos naturales, así la otra orden y concordancia de las otras partes la oración es diversa en cada lenguaje [...].* (Nebrija 1980 [1492] : 204)

‘Ce concert des parties de la phrase entre elles est naturel pour toutes les nations qui parlent, car toutes accordent l’adjectif avec le nom, et le nominatif avec le verbe, et le relatif avec l’antécédent ; mais tout comme ces préceptes sont naturels pour tous, l’autre ordre et la concordance des autres parties de la phrase sont également divers dans chaque langue [...]’

*Entre algunas partes de la oración ai cierta orden casi natural et mui conforme a la razón, en la cual las cosas que por naturaleza son primeras o de maior dignidad, se an de anteponer a las siguientes et menos dignas ; i por esto dize Quintiliano que diremos de oriente a occidente, et no, por el contrario, de occidente a oriente, por que, según orden natural, primero es oriente que el occidente ; et así diremos por consiguiente : el cielo et la tierra, el día et la noche, la luz et las tiniebras, et no por el contrario, la tierra et el cielo, la noche et el día, las tiniebras et la luz.* (Nebrija 1980 [1492] : 205-206)

‘Dans certaines parties du discours, il y a un certain ordre presque naturel et très conforme à la raison, dans lequel les choses qui, par nature, sont les premières ou les plus dignes, doivent être placées avant les suivantes et les moins dignes ; et c’est pourquoi Quintilien dit que nous dirons *d’est en ouest*, et non, au contraire, *d’ouest en est*, parce que, selon l’ordre naturel, l’est est avant l’ouest ; et ainsi nous dirons par conséquence : *ciel et terre, jour et nuit, lumière et ténèbres*, et non, au contraire, *terre et ciel, nuit et jour, ténèbres et lumière*.’

Dans la discussion des violations de l’ordre naturel, Nebrija ne mentionne comme inacceptables que celles qui ne respectent pas l’ordre social. Ainsi le roi doit-il toujours être mentionné en premier et une erreur d’accord comme *vos sois bueno* devrait être plus acceptable que *vuestra merced es bueno* (Nebrija 1980 [1492] : 205-206). Le concept d’ordre naturel chez Nebrija est déterminé par la relation directe que la langue entretient avec le réel et il ne prend pas en compte les règles de construction des phrases.

À partir du XVIII<sup>e</sup> siècle, le conflit entre les rationalistes et les sensualistes, déjà décrit pour la France, se reproduit en Espagne. La grammaire de la Real Academia Española part d’un ordre naturel fondé sur le monde réel extralinguistique. L’Académie répète les exemples

de Nebrija, mais elle mentionne aussi des règles pour l'arrangement des parties du discours : le nom doit précéder l'adjectif parce qu'il désigne la substance et l'adjectif la qualité, le nom substantif doit précéder le verbe parce qu'on doit nommer l'agent avant l'action et l'adverbe doit être postposé parce qu'il ne fait que modifier le verbe :

*Hay un orden natural de colocar las palabras que se funda en la naturaleza misma de las cosas. Este orden pide que el nombre sustantivo prefiera al adjetivo, porque antes es la sustancia que la calidad : que quando el nombre representa al sugeto ó persona que hace, prefiera al verbo, porque ántes es el agente que la acción. Pide que el verbo prefiera al nombre, quando este es el objeto, ó término de la acción ; y al adverbio que califica ó modifica la significación del verbo. (RAE 1771 : 232-233)*

'Il y a un ordre naturel de placement des mots qui est fondé sur la nature même des choses. Cet ordre exige que le nom précède l'adjectif, parce qu'il est la substance avant la qualité : que lorsque le nom représente le sujet ou la personne qui le fait, il précède le verbe, parce qu'il est l'agent avant l'action. Il demande au verbe de précéder le nom, lorsque celui-ci est l'objet ou le but de l'action ; et l'adverbe qui qualifie ou modifie le sens du verbe.'

Cet ordre naturel est bien loin de la réalité linguistique. L'académie remarque qu'il se trouve souvent perturbé pour atteindre des effets stylistiques :

*Pero este orden natural se perturba ó invierte muchas veces para mayor suavidad, elegancia, ó viveza de expresion, según lo pide el asunto de que se trata, y la situacion de las personas que hablan. Menos perturbacion de orden se hallará en las palabras del que habla en una conversacion tranquila, que en la del que está agitado de alguna pasion vehemente: menos por consecüencia en el estilo familiar y didáctico, que en el oratorio, y poético. Llámase construcción natural la que observa con exactitud este orden, y construcción figurada la que no le observa. la construcción natural consta de régimen y concordancia. (RAE 1771 : 234)*

'Mais cet ordre naturel est souvent perturbé ou inversé pour plus de douceur, d'élégance ou de vivacité d'expression, selon les besoins de la matière traitée et la situation des personnes qui parlent. Les paroles du locuteur dans une conversation calme seront moins troublées que celles de l'orateur qui est agité par une passion véhémente : moins dans le style familial et didactique, que dans l'oratoire, et poétique. La construction qui respecte exactement cet ordre s'appelle *construction naturelle*, et *construction figurée* celle qui ne le respecte pas.'

Ces déviations de l'ordre naturel sont décrites comme constructions figurées (*construcciones figuradas*) et elles sont déclarées comme justifiées pour atteindre plus d'élégance et d'énergie. Mais ces constructions figurées vont à l'encontre des règles de la nature. Ce sont des constructions figées qui s'utilisent en lieu et place des constructions naturelles et véritables :

*CONSTRUCCIÓN FIGURADA es la que se aparta de la natural, quando lo piden así el uso ó la mayor elegancia y energia de la expresion. Figura en su recto significado no es otra cosa que ficcion, y en este sentido se usa en la Gramática, porque las expresiones figuradas o fingidas se ponen para sustituir a otras naturales y verdaderas. Por medio de las figuras se altera con frecuencia el ordeng y colocación natural de las palabras : se callan unas : se aumentan otras ; y se falta a las leyes de la concordancia. Quando se invierte el órden se comete la figura hipébaton que vale inversión. Quando se callan palabras es por la figura elipsis que vale falta, efecto. (RAE 1771 : 326-327)*

'La CONSTRUCTION FIGURÉE est celle qui s'éloigne du naturel, quand l'usage ou la plus grande élégance et énergie de l'expression le demandent. La figure dans son sens correct n'est rien d'autre qu'une fiction, et dans ce sens elle est utilisée dans la Grammaire, parce que les expressions figurées ou feintes sont mises à la place d'autres naturelles et vraies. Au moyen des figures, l'ordre et le placement naturels des mots sont souvent modifiés : certains sont réduits au silence, d'autres sont augmentés, et les lois de la concordance font défaut. Lorsque l'ordre est inversé, la figure d'hyperbate est engagée, ce qui veut dire l'inversion. Quand les mots sont réduits au silence, c'est à cause de la figure en ellipse qui veut dire le manque, le défaut.'

La grammaire de l'Académie conçoit l'hyperbate et l'ellipse comme des constructions déficitaires, inévitables lors de certaines utilisations de la langue dans des situations troublées par des émotions, mais elle ne les admet qu'à regret du fait de son attitude normative. L'exemple d'une construction naturelle est décrit comme correspondant aux règles :

*El premio y el castigo son convenientes en la guerra, así como la justicia y la clemencia son convenientes en la paz. (RAE 1771 : 327)*

'La récompense et le châtement sont convenables en temps de guerre, tout comme la justice et la miséricorde sont convenables en temps de paix.'

En revanche, dans la construction suivante qui est figurée, l'ordre naturel des mots a été modifié par le déplacement de l'adverbe avant le verbe et de l'adjectif avant les noms. De plus, l'omission de l'adjectif *convenientes* entraîne la structure déficiente de l'ellipse :

Así como son convenientes en la paz la justicia y la clemencia, son en la guerra el premio, y el castigo. (RAE 1771 : 328-329)

‘Tout comme la justice et la miséricorde sont commodes dans la paix, la récompense et le châtement le sont dans la guerre.’

Bien que l'inversion soit considérée comme une perturbation de l'ordre naturel (*una perturbacion del orden natural*, RAE 1771 : 329) l'attitude normative de l'Académie n'est pas stricte au point de la rejeter entièrement. Elle considère même qu'il est parfois approprié de changer l'ordre naturel, ce que confirment des exemples linguistiques. Même si l'ordre figuré peut sembler arbitraire, il est fondé sur la convenance (*conveniencia*) de l'expression selon la raison et l'autorité ou, tout au moins, selon l'autorité des modèles linguistiques, si aucune raison ne peut être trouvée (RAE 1771 : 333). L'attitude pragmatique de l'Académie dans la normalisation de l'ordre des mots conduit à un brouillage de la frontière entre le naturel et le figuré :

*El uso es tan poderoso que ha hecho ya cómo naturales y comunes muchas expresiones figuradas, de tal suerte que serían defectuosas si se quisiese reducirlas al riguroso orden natural.* (RAE 1771 : 333)

‘L'usage est si puissant qu'il a déjà rendu de nombreuses expressions figurées naturelles et banales, de sorte qu'elles seraient défectueuses si elles devaient être réduites à l'ordre naturel strict.’

L'attitude de l'Académie royale espagnole à l'égard de la normalisation de l'ordre des mots est donc assez libérale. Tout en déclarant l'inversion comme une déviation par rapport aux règles, elle l'accepte comme naturelle pour atteindre certains buts communicatifs.

Certains auteurs espagnols défendent les mêmes idées. Par exemple, Antonio Capmany i Suris de Montpalau (1742-1813) consacre une section de son *Arte de traducir* à l'inversion ou transposition. Il aborde ce problème à partir de la théorie de l'ordre naturel rationaliste, qu'il considère comme incontestable, et d'un ensemble arbitraire qui dépend du génie de la langue :

*Esta colocación de las palabras, independiente del orden y sucesion natural de las ideas, parece que ha sido arbitraria en todas las lenguas, y mas en la latina que gozaba de todas las licencias imaginables de la transposicion. Las lenguas vivas han adoptado en esta parte su método, que llamaremos construccion usual. El modismo de la francesa es mas riguroso y atado ; pero la española se sujeta menos à un orden fixo y monótono.* (Capmany 1987 [1776] : 138)

‘Ce placement des mots, indépendant de l’ordre naturel et de la succession des idées, semble avoir été arbitraire dans toutes les langues, et plus encore en latin, qui a bénéficié de toutes les licences de transposition imaginables. Les langues vivantes ont adopté dans cette partie leur méthode que nous appellerons la construction habituelle. L’idiome de la langue française est plus rigoureux et plus lié ; mais la langue espagnole est moins soumise à un ordre fixe et monotone.’

Capmany ne lie pas les libertés dont jouit la langue latine à l’existence de cas, marqueurs morphologiques des rapports entre les parties du discours. Il ne traite pas non plus des raisons qui rendent l’ordre des mots en français plus rigide qu’en espagnol. Si la langue espagnole est moins sujette à un ordre monotone et fixe (« *sujeta menos à un orden fixo y monótono* », Capmany 1987 [1776] : 138), cela peut être vu comme un avantage : elle échappe à la monotonie.

Pour comparer l’ordre des mots en français et en espagnol, Capmany donne 70 exemples (Capmany 1987 [1776] : 138-141) de phrases françaises traduites en espagnol, parmi lesquelles se trouvent des séquences qui usent de structures syntaxiques différentes pour marquer la même structure informationnelle. Ainsi :

- |  |  |
|--|--|
| a. Il a tout donné.                                  | Todo lo ha dado.   |
| b. Tant il était ignorant.                           | Tan ignorante era.   |
| c. Occupés du nécessaire<br>les sauvages ne pensent. | Los Salvages, ocupados en buscar<br>el sustento, no piensan. |
| d. Que vous êtes importun.                           | Qué importuno sois.  |

Le Galicien éclairé Luis Marcelino Pereira (1754-1811) s’est également exprimé contre l’hypothèse rationaliste d’un ordre naturel des idées. L’ordre dans lequel les idées se présentent à l’esprit est celui que l’on pourrait le plus justement qualifier de naturel, ou « celui de leur hiérarchie, je veux dire celui de leur plus ou moins grande importance dans le discours » (Pereira [entre 1798 et 1800]), celui de la plus ou moins grande attention accordée à chaque idée par l’orateur

ou l'auditeur. La discussion sur l'ordre des mots est liée à l'efficacité de la communication, qui est décrite comme la fonction principale du langage. Pereira contredit l'idée d'une relation obligatoire entre la dépendance et la place qu'un mot occupe dans la phrase. L'ordre des mots est seulement un des moyens de subordination :

*De manera que muy lexos de guardarse en Darium vicit Alexander, como en la otra construccion, el orden de la dependencia de la ideas, se sigue otro que le es de todo punto contrario. Lo que si digo es que aunque mas sea, como lo es en verdad, necesario dar á entender en la oracion clara y distan-tamente el modo como se lian entre si todas las ideas, y las unas dependen de las otras ; ninguna necesidad hay de que se haga esto señalando lugar determinado á todas las palabras. Pues ni el orden de estas tiene por natu-raleza con la subordinacion, ó dependencia de aquellas relacion alguna : ni es mas sino un medio de que apropósito de declararla podemos como de otros aprovecharnos, tan solo necesarios en falta de ellos. (Pereira [entre 1798 y 1800], f. 79 )*

‘Ainsi, comme dans l’autre construction, l’ordre de dépendance des idées est très différent de celui de *Darium vicit Alexander*. Ce que je dis, c’est que même s’il est nécessaire, comme c’est le cas en vérité, de préciser dans la phrase la manière dont toutes les idées sont liées entre elles, dont l’une dépend de l’autre, il n’est pas nécessaire de le faire en indiquant une certaine place à toutes les paroles. Car l’ordre de ces dernières n’est ni subordonné par nature à ces relations, ni dépendant de celles-ci ; il n’est pas non plus un moyen par lequel nous pouvons, pour les déclarer, en tirer profit, comme nous pouvons le faire pour les autres, il est seulement nécessaire en leur absence.’

En raison de l’importance du thème du discours en France, la question de l’ordre des mots gagne inévitablement d’autres pays. On peut observer ici une transmission linéaire de la doctrine de l’ordre naturel, qui conduit à la demande d’adapter l’ordre des mots d’autres langues à l’ordre fixe des mots sujet-*verbe*-*objet*, y compris pour des langues à cas comme l’allemand ou le russe. Selon Carl Friedrich Aichinger (1717-1782), l’allemand doit envisager un ordre fixe des mots :

*Die Lateiner und Griechen können ihre Theile der Rede in unterschiedliche Ordnung stellen, ohne, daß der Verstand darunter Schaden leidet, z. B. bellum dulce est inexpertis, inexpertis bellum est dulce, dulce est bellum inexpertis, est dulce bellum inexpertis, &c. Die Teutschen aber haben eine*

*vest gestellte Ordnung in welcher ihre Wörter stehen müssen.* (Aichinger 1754 : 525-526)

‘Le latin et le grec peuvent mettre leurs parties de discours dans un ordre différent sans nuire à l’esprit, par exemple *bellum dulce est inexpertis, inexpertis bellum est dulce, dulce est bellum inexpertis, est dulce bellum inexpertis*, etc. Les Allemands, cependant, ont un ordre fixe dans lequel leurs mots doivent être placés.’

Johann Christoph Gottsched (1700-1766) met également en perspective l’hypothèse d’un ordre naturel des mots. Il pense qu’il est naturel que tous les peuples considèrent l’ordre des mots dans leurs langues respectives comme correspondant le mieux à la pensée logique. Selon Gottsched, ils ne voient pas que ce sont les langues elles-mêmes qui leur ont enseigné ce genre de pensée :

*In einer allgemeinen Sprachlehre könnte man zeigen, welches die natürlichste Ordnung der Gedanken wäre, die in einer philosophischen Sprache beobachtet werden müßte. Hier ist es genug zu bemerken, daß fast jedes Volk sich einbildet, seine Art die Wörter zu setzen, sey der Natur der Gedanken die gemäßigste. Allein sie irren alle ; und bemerken nicht : daß ihnen ihre Art zu denken, zuerst durch ihre Muttersprache beygebracht worden.* (Gottsched 1762 [1748] : 399)

‘Dans une théorie générale du langage, on pourrait montrer quel serait l’ordre le plus naturel des pensées à considérer dans un langage philosophique. Il suffit ici de constater que presque chaque peuple imagine que sa propre façon de disposer les mots est la plus appropriée à la nature de la pensée. Mais ils sont tous dans l’erreur, et ils ne se rendent pas compte que leur façon de penser leur a été enseignée d’abord par leur langue maternelle.’

Michail Vasil’evič Lomonosov (1711-1765) aussi est en faveur de l’idée de l’ordre naturel déterminé rationnellement et contre un ordre de mots « artificiel », avec des inversions ; mais il préconise des inversions pour des raisons esthétiques (Breuillard 2010 : 48). Nikolaj Karamzin (1766-1826) est favorable à l’idée de l’ordre libre des mots, affirmant que chaque changement dans une séquence de mots entraîne un changement de contenu et donne à l’expression un sens particulier. Selon lui, il est erroné de dire *zemlju plodotvorit solnce* ou *plodotvorit solnce zemlju* en lieu et place de *solnce plodotvorit zemlju* (« le soleil féconde la terre »). La grammaire russe ne spécifie pas l’ordre des mots (Breuillard 2010 : 52). Karamzin critique Lomonosov :

ce dernier ne suit pas toujours ses pensées dans l'ordre des mots et utilise des constructions qui ne sont pas toujours agréables à l'oreille.

L'accent mis sur la pauvreté flexionnelle des autres langues conduit à un discours encore plus riche sur ce problème et finalement à l'idée que l'ordre des mots est un des moyens d'exprimer les relations communes dans la phrase, qui interagissent dans cette fonction dans des proportions différentes.

En accord avec ce qui précède, Joseph Priestley (1733-1804) caractérise l'ordre dans lequel les mots sont positionnés comme non déterminant tant que leurs relations sont rendues suffisamment explicites par des marques flexionnelles. L'inversion du sujet et du prédicat est même possible en anglais, malgré sa pauvreté flexionnelle, si les conditions correspondantes sont remplies. Ainsi, dans les séquences de mots suivantes, il est parfaitement clair que seul Jésus peut être l'acteur, sa position à la fin de la phrase est donc possible : *These things did Jesus. Jesus did these things* (Priestley 1762 : 255-256).

Même lorsque l'inversion du sujet et de l'objet n'est pas possible, il existe des constructions qui permettent d'établir un ordre correspondant à des fins communicatives, comme par exemple la passivation. Dans l'ensemble, l'ordre des mots est réglé non pas tant par la nécessité absolue que par l'attention de l'orateur à certaines parties de l'énoncé, l'harmonie ou l'adaptation au familier :

*But not only many a writer take the liberty to change the order of his words in a sentence ; but, in most languages, every complex sentence is capable of a great variety of expression, by little more than a different construction of the same words ; which contribute greatly to the ease and harmony of composition. I shall instance in the last mentioned sentence the sense of which may be expressed by saying either, that Alexander conquered Darius, or Darius was conquered by Alexander. (Priestley 1762 : 161-163)*

‘Non seulement beaucoup d'écrivains prennent la liberté de changer l'ordre de leurs mots dans une phrase, mais, dans la plupart des langues, chaque phrase complexe est capable d'une grande variété d'expressions, en utilisant à peine plus qu'une construction différente des mêmes mots, ce qui contribue grandement à la facilité et à l'harmonie de la composition. Je citerai en exemple la dernière phrase mentionnée, dont le sens peut être exprimé en disant soit qu'*Alexander conquered*

*Darius* (« Alexandre a conquis Darius »), soit que *Darius was conquered by Alexander* (« Darius a été conquis par Alexandre »).’

L’intensité de la discussion menée en France autour de l’ordre des mots donne donc lieu à une tentative d’adaptation de la séquence sujet-verbe-objet à d’autres langues, à partir de positions rationalistes. L’objectif est d’obtenir une compréhension plus simple des énoncés, puisque l’expression linguistique suit alors l’ordre de la pensée. Cependant, l’existence d’autres moyens linguistiques pour marquer la structure de l’information (comme les flexions nominales), mais également la prise en compte accrue des besoins de la rhétorique et le développement de théories linguistiques sensualistes limitent ces efforts normatifs.

#### 4. Continuité et conclusions

Le concept d’ « ordre naturel », issu de la rhétorique et associé au départ à des grammaires normatives, se développe, en France, sous l’effet de principes philosophiques rationalistes et devient un élément important de la clarté de la langue française. Avec l’émergence des théories linguistiques sensualistes et la prise en compte accrue des aspects pragmatiques de l’usage réel de la langue, le concept d’ordre naturel revêt un autre contenu et se mesure alors à l’intention du locuteur. Avec un certain retard, des discussions similaires ont lieu dans d’autres pays ; leur portée est néanmoins atténuée : il s’agit d’autres traditions grammaticales et d’autres conditions linguistiques.

L’hypothèse d’un ordre naturel qui permettrait d’économiser des « coûts » cognitifs connaît également une étonnante continuité aux XIX<sup>e</sup> et XX<sup>e</sup> siècles. Henri Weill (1991 [1844]) aborde le sujet et en précise plusieurs aspects : les langues anciennes ou modernes, la maîtrise des langues avec un ordre fixe des mots ou le cas des langues avec un ordre libre de mots, la logique grammaticale ou rhétorique. Georg von der Gabelentz (1840-1893) (1901 [1891]) et Hermann Paul (1846-1921) (1880) déclarent que le sujet psychologique est celui vers lequel l’orateur veut diriger l’attention directe du récepteur et que le prédicat psychologique est ce que le récepteur doit penser du sujet. De ce point de vue, le sujet peut être différencié, mais le prédicat grammatical doit néanmoins rester ancré sur des fondements

psychologiques, car ces derniers reflètent les distributions typiques de la structure informationnelle. Paul souligne également la diversité des moyens d'expression, en mentionnant à la fois « l'intensité du son » et l'ordre des mots.

Le fondateur de la philologie romane, Friedrich Diez (1794-1876), mentionne déjà la possibilité de mettre des éléments non sujets au début de la phrase ainsi que l'existence de phrases clivées. Il utilise les termes très courants à l'époque de *Nachdruck* ('insistance') et *Hervorhebung* ('mise en relief') pour décrire ces phénomènes. L'auteur de la plus éminente grammaire romane du XIX<sup>e</sup> siècle, Wilhelm Meyer Lübke (1861-1936), conçoit l'ordre des mots comme une « lutte » (*Kampf*) entre les facteurs grammaticaux et affectifs. En lien avec ces derniers, il observe une tendance à mettre l'élément le plus important ou remarquable à la fin de la phrase (Dufter & Jacob 2007 : 1).

Dans les travaux de recherche récents sur la structure informationnelle, l'hypothèse d'une transformation de la phrase structurelle (naturelle) en un ordre linéaire ou d'une augmentation des coûts cognitifs du fait de l'utilisation d'une forme inversée et marquée peut être considérée comme représentative de la réflexion au XX<sup>e</sup> siècle (cf. Enkvist 1986, Krifka & Féry 2008), par exemple :

*German IPs have a "normal" word order (Lenerz 1977), and deviations from this normal order serve the expression of a marked topic-focus structure. (Fanselow 2003 : 197)*

'Les IP [*inflection phrases* « phrases de flexion »] allemands ont un ordre de mots « normal » (Lenerz 1977), et les déviations de cet ordre normal servent à l'expression d'une structure topic-focus marquée.'

La détermination de cet ordre « normal » paraît possible d'un point de vue logique, mais d'un point de vue pragmatique, il est tout à fait incertain que cet ordre cause moins d'efforts cognitifs qu'un ordre différent dérivé de l'intention communicative du locuteur.

## Bibliographie

- Aichinger, Carl Friedrich. 1754. *Versuch einer teutschen Sprachlehre, anfänglich nur zu eignem Gebrauche unternommen, endlich aber, um den Gelehrten zu fernerer Untersuchung Anlaß zu geben, ans Liecht gestellt von Carl Friedrich Aichinger*. Franckfurt, Leipzig : Johann Paul Kraus.
- Arnauld, Antoine, Claude Lancelot. 1768 [1660]. *Grammaire générale et raisonnée : contenant les fondemens de l'art de parler [...] les raisons de ce qui est commun à toutes les langues [...]*, Éd. par l'Abbé Fromant. Paris : Prault Pere.
- Batteux, Charles. 1747-1748. *Cours de belles-lettres distribué par exercices*. Paris : Desaint et Saillant.
- Batteux, Charles. 1763. *De la Construction oratoire*. Paris : Desaint & Saillant.
- Bouhours, Dominique. 1671. *Les Entretiens d'Ariste et d'Eugène*. Paris : Sebastien Mabre-Cramoisy.
- Breuillard, Jean. 2010. L'ordre des mots dans la pensée grammaticale russe du 18<sup>e</sup> siècle. *Histoire Épistémologie Langage* 32(1). 47-68.
- Capmany, Antonio. 1987 [1776]. *Arte de traducir el idioma francés al castellano con vocabulario lógico y figurado de la frase comparada de ambas lenguas*. Éd. par María del Carmen Fernández Díaz. Santiago de Compostela: Universidad de Santiago de Compostela.
- Charpentier, François. 1683. *De l'excellence de la langue française par M. Charpentier, de l'Académie Française*. Paris : Veuve Bilaine.
- Chevalier, Jean-Claude. 2006 [1968]. *Histoire de la syntaxe. La naissance de la notion de complément dans la grammaire française (1530-1750)*. Paris : Éditions Champion.
- Condillac, Étienne Bonnot de. 1961 [1746]. *Essai sur l'origine des connoissances humaines : ouvrage où l'on réduit à un seul principe tout ce qui concerne l'entendement humain*, Reprod. de l'éd. de 1746. Paris : INALF.
- Cordemoy, Géraud de. 1668. *Discours physique de la parole*. Paris : F. Lambert.
- Delesalle, Simone. 1980. L'évolution de la problématique de l'ordre des mots du 17<sup>e</sup> au 19<sup>e</sup> siècle en France. L'importance de l'enjeu. *Des ordres en linguistique* 22-23. 235-278.
- Diderot, Denis. 1969 [1751]. Lettre sur les sourds et muets. *Œuvres complètes de Denis Diderot, édition chronologique, introduction de Roger Lewinter*, vol. II. Paris : Société Encyclopédique Française et le Club Français du Livre. 513–602.
- Du Marsais, César Chesneau. 1797. *Œuvres*. Éd. de Marie-Émile-Guillaume Duchosal & Charles Million. Paris : De l'imprimerie de Pougin.
- Dufter, Andreas & Daniel Jacob, éd. 2007. *Focus and Background in Romance Languages*. Amsterdam: John Benjamins Publishing Company.

- Enkvist, Nils Erik. 1986. Linearization, text type, and parameter weighting. *Language and Discourse: Test and Protest. A Festschrift for Petr Sgall*, dir. par Jacob L. Mey. Amsterdam: John Benjamins Publishing Company. 245-260.
- Fanselow, Gisbert. 2003. Free constituent order: a minimalist interface account. *Folia Linguistica* 37. 191-232.
- Gabelentz, Georg von der. 1901 [1891]. *Die Sprachwissenschaft : ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Leipzig : Tauchnitz.
- Gottsched, Johann Christoph. 1762 [1748]. *Vollständigere und Neuerläuterte Deutsche Sprachkunst. Nach den Mustern der besten Schriftsteller des vorigen und itzigen Jahrhunderts abgefasst, und bey dieser fünften Auflage merklich verbessert von Johann Christoph Gottscheden*. Leipzig : Breitkopf und Sohn.
- Haßler, Gerda. 2012. La linéarité du langage comme problème théorique dans les théories rationalistes et sensualistes. *Beiträge zur Geschichte der Sprachwissenschaft* 22(1). 33-66.
- Haßler, Gerda. 2014. "Inversions of word order generate higher costs". Continuity and development of a topos since the rationalist language theories of the 17<sup>th</sup> century. *History of Linguistics 2011. Selected papers from the 12<sup>th</sup> International Conference on the History of the Language Sciences (ICHOLS XII)*, dir. par Vadim Kasevich, Yuri Kleiner, Patrick Sériot. Amsterdam, Philadelphie : John Benjamins Publishing Company. 3-11.
- Jouvancy, Joseph de. 1692. *Oratio habita in regio Ludovici Magni collegio*. Paris: F. Muguet.
- Krifka, Manfred & Caroline Féry. 2008. Information structure. Notional distinctions, ways of expression. *Unity and Diversity of Languages*, dir. par Piet van Sterkenburg. Amsterdam : John Benjamins Publishing Company. 123-136.
- Lamy, Bernard. 1688 [1675], *La Rhétorique ou l'art de parler*. Troisième édition revue & augmentée. Paris : André Pralard.
- Lancelot, Claude. 1650 [1644]. *Nouvelle méthode pour apprendre [...] la langue latine*. Seconde édition, revue et augmentée [...] avec une brève instruction sur les règles de la poésie françoise. Paris : A. Vitré.
- Lancelot, Claude. 1656 [1655]. *Novvelle methode pour apprendre facilement la langue grecque, contenant les regles des declinaisons, des conivugaisons de l'Inuestigation du Thème, de la Syntaxe, de la Quantité des Accens, des Dialectes, & des Licences Poétiques ; mis en François dans un ordre très clair & très abrégé. Avec un grand nombre de Remarques très solides & très nécessaires pour la parfaite connoissance de cette Langue, & pour l'intelligence de ses Auteurs*. Seconde édition. Paris : Pierre Le Petit.

- Le Laboureur, Louis. 1669. *Avantages de la langue françoise sur la langue latine*. Paris : G. de Luyne.
- Lenerz, Jürgen. 1977. *Zur Abfolge nominaler Satzglieder im Deutschen*. Tübingen : Narr.
- Maupas, Charles. 1632 [1618]. *Grammaire et syntaxe françoise, contenant reigles bien exactes & certaines de la prononciation, orthographe, construction & usage de nostre langue, en faveur des estrangiers qui en sont desireux*. Troisième édition, reveuë, corrigée & augmentée de moitié, & en beaucoup de sortes amendée, entre la precedente, par le dit Auteur. Roven : Jacques Cailiove.
- Meigret, Louis. 1550. *Le tretté de la grammaire françoeeze*. Paris : Chrestien Wechel.
- Nebrija, Antonio. 1980 [1492]. *Gramática de la lengua castellana*. Estudio y edición de Antonio Quilis. Madrid : Editora nacional.
- Paul, Hermann. 1880. *Principien der Sprachgeschichte*. Halle : Niemeyer.
- Pereira, Luis Marcelino. [entre 1798 et 1800]. *Discurso sobre las mejoras y menoscabos que recibió la lengua Castellana en los siglos XVII y XVIII*, RAE manuscrit MS 112. Real Academia Española.
- Priestley, Joseph. 1762. *A Course of Lectures on the Theory of Language and Universal grammar*. Warrington : W. Eyres.
- RAE. 1771. *Gramática de la lengua castellana*. Compuesta por la Real Academia Española. Madrid : Ibarra.
- Ricken, Ulrich. 1978. *Grammaire et philosophie au siècle des Lumières : controverses sur l'ordre naturel et la clarté du français*. Villeneuve-d'Ascq : Publications de l'Université de Lille III.
- Ricken, Ulrich. 1994. *Linguistics, Anthropology and Philosophy in the French Enlightenment : Language Theory and Ideology*. Translated from the German by Robert E. Norton. London, New York : Routledge.
- Seguin, Jean. 1993. *L'invention de la phrase au XVIII<sup>e</sup> siècle : contribution à l'histoire du sentiment linguistique français*. Louvain : Peeters et Paris : Société pour l'information grammaticale.
- Weill, Henri. 1991 [1844]. *De l'ordre des mots dans les langues anciennes comparées aux langues modernes. Question de grammaire générale*. Paris : Didier Érudition.

# Pourquoi l'italien serait-il la « meilleure » langue pour chanter ? ou : Comment la sonorité peut déterminer l'impression de simplicité d'une langue

**Claudia Schweitzer**

Praxiling, Université Paul Valéry Montpellier 3, CNRS, UMR 5267 / Histoire des théories linguistiques, UMR 7597.

---

## RÉSUMÉ

Dans sa *Lettre sur la musique française* (1753), Rousseau déclare « qu'il n'y a ni mesure ni mélodie dans la musique française, parce que la langue n'en est pas susceptible ». Rousseau, comme beaucoup d'auteurs jusqu'à aujourd'hui, considère le français comme une langue inadaptée pour chanter, tandis que l'italien s'y prête particulièrement bien, grâce à sa richesse en voyelles qui lui assurent une grande sonorité. Il semble logique que plus une langue est riche en voyelles, plus elle est chantante. Comme à l'époque baroque la musique est modelée sur le rythme, la hauteur et la couleur de chaque langue, une base langagière très sonore donnera aussi un chant plus sonore. Mais aujourd'hui, on oublie souvent que les esthétiques musicales italienne et française sont très différentes à l'âge baroque. Si la musique italienne se veut sonnante et sensuelle, la française donne l'importance à la compréhension du texte, assurée par la bonne compréhension des consonnes. Et chez les auteurs du XVIII<sup>e</sup> siècle, un deuxième sujet se joint également à ce type de discussion : le caractère ou énergie de la langue qui reflète celui de la nation qui la parle. Avec ces réflexions, on rejoint celles sur le « génie de la langue ».

Cet article propose un regard sur les caractéristiques d'une langue qui

Schweitzer, Claudia. 2023. Pourquoi l'italien serait-il la « meilleure » langue pour chanter ? ou : Comment la sonorité peut déterminer l'impression de simplicité d'une langue. *Simplicité et complexité des langues dans l'histoire des théories linguistiques*, dir. par Chloé Laplantine, John E. Joseph & Émilie Aussant. Paris: SHESL (HEL Livres, 3). 55-78.

font qu'elle est considérée comme bonne ou simple pour le chant. A partir des définitions du mot *simple*, données dans *Le Petit Robert*, nous proposons une réflexion autour de trois synonymes ou paraphrases d'une *langue simple* : une langue *naturelle*, une langue *facile et commode* et une langue *évidente*.

#### MOTS-CLÉS

simplicité et complexité des langues, italien, français, chant baroque, caractères des peuples

#### ABSTRACT

In his *Lettre sur la musique française* (1753), Rousseau declared "that there is neither measure nor melody in French music, because the language is not capable of it". Rousseau, like many up to the present day, considers French a language unsuitable for singing, while Italian lends itself particularly well to song, thanks to its richness in vowels which give it a great sonority. It seems logical that the richer a language is in vowels, the more melodious it is. Since in the Baroque period the music was modeled on the rhythm, pitch and color of each language, a very sonorous language base would result in a more sonorous song. But today, we often forget that Italian and French musical aesthetics were very different in the Baroque age. If Italian music aims to be resonant and sensual, French music emphasizes understanding of the text, ensured by good understanding of the consonants. And with eighteenth-century authors, a second subject figures as well in discussions of this sort: the character or energy of the language, which reflects that of the nation that speaks it. With these considerations, we link up with the discourse on the «genius of the language».

This article offers a look at the characteristics of a language which lead to it being considered good or simple for singing. Starting from the definitions of the word *simple* given in *Le Petit Robert*, we explore three synonyms or paraphrases of a *simple* language: a *natural* language, an *easy and convenient* language and an *obvious* language.

#### KEYWORDS

simplicity and complexity of languages, Italian, French, baroque singing, characters of peoples

## 1. Introduction

C'est aujourd'hui un lieu commun de penser que l'italien est la langue de référence de la musique. On dit que l'italien se chante bien

et que cette langue est tellement mélodieuse que, même parlée, elle semble déjà chanter. Un bref regard dans la presse, plus ou moins populaire ou scientifique, confirme l'omniprésence de cette opinion qui se manifeste soit par une louange directe, soit par la dénégation d'une autre langue :

La langue italienne, une langue musicale. [...] L'italien [...] est bel et bien le langage de la musique ! [www.superprof.fr, consulté le 10/01/2020]

L'italien est la langue reine dans le monde de la musique et de l'Opéra. [www.scuola-toscana.com, consulté le 10/01/2020]

La langue française n'est pas musicale. (Bernard Cerquiglini. *Le Figaro* 22/6/2018)

En italien, je ne dois produire aucun effort, je me régale à dire les mots et tout devient plus facile, même si la couleur exacte des voyelles est difficile à trouver (Natalie Dessay, Entretien, *Opéra magazine* 2005(2) : 8)

La langue italienne est [...] souvent chérie par l'expérience proprioceptive de nombreux interprètes du répertoire vocal. (Paolo Zedda. « Langues chantées ». *La Revue du Conservatoire*. 2013)

Langue de prédilection du répertoire de l'opéra, l'italien occupe jusqu'à aujourd'hui – et cela, malgré la domination mondiale de la langue anglaise – le statut de *langue de la musique (classique)*. L'emploi d'un grand nombre de termes techniques empruntés à l'italien (*soprano, basso, crescendo, diminuendo, da capo al fine, legato...*) semble confirmer cette position de référence.

Historiquement, le rôle prédominant de l'italien commence à l'époque baroque où la musique italienne est déjà tellement omniprésente dans le monde musical que Nicolaus Harnoncourt (1984 : 199) note que l'on « perçoit la musique française de cette époque presque comme une réaction à cette éruption musicale ». La question de la langue de la musique et du chant est aussi au cœur de la Querelle des Bouffons (1752-1754), marquée par le fameux dénigrement de la musique française par Jean-Jacques Rousseau, selon qui la musique de son peuple n'a « ni mesure ni mélodie », et cela pour la simple raison que « la langue n'en est pas susceptible » (Rousseau 1753). Effectivement, selon l'opinion de Giambattista Mancini,

Toutes les nations sont obligées, bon gré ou malgré [sic !], de convenir que la langue italienne est, de toutes les langues, la plus harmonieuse,

la plus douce, la plus sonore, la plus propre, en un mot, à être adoptée à une bonne musique. (Mancini 1776 : 199)

On peut alors se demander quelles sont les raisons pour ce grand succès de l'italien dans le monde musical ? Comme le remarque Paolo Zedda (2013), beaucoup de constats favorables à la langue italienne sont liés à la proprioception des artistes, elles renvoient alors à des critères techniques. L'argument de Rousseau en revanche est entièrement d'ordre esthétique. Selon lui, la langue italienne est plus riche en voyelles et par conséquent mieux adaptée au chant. Pour la même raison, elle est plus expressive que la langue française (ce qui suppose que l'on soit d'avis que les voyelles sont favorables au chant et porteuses de l'expressivité du chant).

Cet article propose un regard sur les caractéristiques qui font qu'une langue est considérée comme bonne ou simple pour le chant. Nous nous intéressons aux caractéristiques sonores des langues – notamment de l'italien en comparaison avec le français – car évidemment, pour chanter, les traits phoniques et sonores d'une langue sont de la plus grande importance. A partir des définitions du mot *simple*, données dans *Le Petit Robert*, nous proposons une réflexion autour de trois synonymes ou paraphrases d'une *langue simple* : une langue *naturelle*, une langue *facile et commode* et une langue *évidente*.

## 2. Une langue naturelle : une question esthétique

D'un point de vue esthétique, une langue naturelle devra être celle qui rend naturellement et sans effort supplémentaire<sup>1</sup> les intentions du locuteur, qu'elles soient raisonnées ou émotives. Quant à la musique, langue des émotions selon la citation populaire, attribuée à Emmanuel Kant (« la musique est la langue des émotions »), il s'agit évidemment avant tout du deuxième critère. Une œuvre musicale transporte un ensemble qui se compose d'un texte (pour le chant) et de la musique, quant à elle expression des émotions ou des passions

---

1. Le *Petit Robert* donne comme une explication du terme *simple* que la chose désignée « comporte peu d'éléments ajoutés ».

humaines : « Il semble que comme la parole est l'art de transmettre les idées, la mélodie soit celui de transmettre les sentiments » (Rousseau 1755 : 337<sup>2</sup>).

A l'âge baroque (musical), texte et musique rentrent dans une union extrêmement étroite. Pour écrire sa partition, le compositeur part des mots, de leurs sonorités, de la quantité de leurs syllabes, des groupes de mots ou phrases, de leur débit, de leurs caractéristiques intonatives et de leurs qualités rythmiques. Ainsi, le texte n'est pas greffé par-dessus, mais sa substance est à l'origine même de la musique. La matérialité de la langue est traduite par les durées, par les accents d'intensité, par les suppressions, les silences, les timbres, etc. Ce concept n'a jamais été théorisé tel quel par les musiciens baroques : il découlait naturellement de la subordination de la musique à la parole et il en allait ainsi « de soi »<sup>3</sup> (cf. Cannone 2006). On discute en revanche sur les qualités et les valeurs esthétiques, ou tout simplement sur la primauté de la musique d'une nation sur celle d'une autre, et cela, en particulier par rapport aux musiques italienne et française. L'allemand Johann Gottlieb Walter par exemple, un cousin de Jean Sébastien Bach, analyse dans l'article « Stylus (lat.), stilo (ital.), stile (gall.) » de son dictionnaire de musique (1732) la différence perceptive, caractérisant le style italien comme incisif, coloré et expressif (« *scharff, bunt und ausdrückend* ») et le style français comme naturel, fluide et tendre (« *natürlich, fließend, zärtlich* »).

A l'âge baroque, il semble évident que la musique doit être différente selon la langue utilisée : français ou italien, mais aussi allemand, anglais... Les écarts mélodiques plus importants dans la musique italienne, les changements de tempo plus fréquents, le

---

2. Les effets de la musique sur les émotions et sur les humeurs de l'auditeur ont été discutés par les chercheurs anciens et modernes, mais on les trouve toujours aussi dans le langage quotidien des gens. Citons à titre d'exemple la lettre de Madame de Sévigné, écrite le 8 janvier 1674 à Madame de Grignon, où elle commente une visite de la tragédie lyrique *Cadmus* de Jean-Baptiste Lully comme suit : « L'Opéra [...] est un prodige de beauté : il y a des endroits de la musique qui m'ont déjà fait pleurer. » (Sévigné, [1806], lettre 280 : 446).

3. En effet, on en trouve les premières réflexions théoriques écrites justement au moment où le concept est remis en cause, suite au changement esthétique musical.

timbre cherché par l'instrumentation et les tessitures préférées plus aiguës et brillantes – tous ces aspects, comme maints d'autres, trouvent leur origine dans ce mode de composition. Les différentes manières de *parler*, c'est-à-dire de produire et de combiner les sons langagiers, ne font que soutenir et renforcer la perception différente des deux styles de musiques<sup>4</sup>.

Ainsi, en italien, les voyelles semblent omniprésentes. Beaucoup de mots finissent avec une sonorité vocalique, même si le système ne connaît qu'un nombre restreint de voyelles (Rogers & d'Arcangeli 2004). Ces voyelles se désignent par une grande netteté, les rendant facilement distinguables et identifiables (Écalle & Magnan 2002 : 102)<sup>5</sup>. Le corps entier du locuteur est investi dans la production des sons. L'air ou le souffle sont traités de manière flexible au niveau des abdominaux, ce qui permet de donner plus d'accentuation (*stress* en anglais) ou de force, et plus d'emphase dynamique aux sons ou, autrement dit, plus d'accent aux syllabes (Garde 1968 ; Canepari 1976). Cet effet concerne notamment les sons vocaliques (Sandford 1995). Généralement, en italien, la syllabe accentuée est également, au moins légèrement, allongée (Roegiest 2009 : 149). La flexibilité de cet accent dont le déplacement peut servir à la compréhension, n'est possible que grâce à la gestion souple de l'air de l'expiration.

En français, le souffle est traité de manière plus statique, les écarts sont moins sensibles et la chaîne parlée semble plus plate. Les voyelles ressortent moins et la différence de leur longueur est de toute façon incomparablement moins développée qu'en italien (Sandford 1995). Voyelles comme consonnes se caractérisent par une attaque douce qui croît lentement (Delattre 1953). La syllabe connaît deux mouvements : le premier est une ouverture progressive et le second une fermeture vive. Initiant et assurant la transition syllabique, ce deuxième mouvement n'influe pas sur la voyelle précédente. Le son des consonnes sonnantes en position finale est maintenu par une

---

4. Sur le site <https://sscm-jscm.org/v1/no1/sanford.html>, on trouve des enregistrements, fait par la chanteuse Sally Sanford, qui mettent bien en évidence les distinctions dont nous parlons dans la suite.

5. Ajoutons que la valeur %V qui mesure la proportion de durée des intervalles vocaliques de l'énoncé de l'italien (4,5) est sensiblement plus élevée que celle du français (3,9). Voir Ramus (1999 : 47).

détente musculaire (Pierret 1994 : 89). Évidemment, cette sonorité n'égale pas celle des voyelles ouvertes italiennes à la fin des mots. L'accent tombant toujours à la fin du mot ou d'un groupe de mots, l'ensemble de la structure orale semble plus rigide par l'accentuation fixe (Saffi 2001).

Aussi, l'Italien Francesco Algarotti parle-t-il en 1750 du français comme une langue (entre autres) peu harmonieuse, peu varié et peu capable de refléter les différents mouvements de l'âme, bref, comme une langue peu musicale<sup>6</sup>. Passant à la voix chantée, le même type de remarque s'impose, au moins si l'on regarde les anciennes descriptions.

La technique des chanteurs baroques français, naturellement adaptée à la langue française, rend l'expressivité du texte par le traitement des sons consonantiques. La bonne prononciation des consonnes et leur sonorisation prennent une place considérable dans les grands traités du XVIII<sup>e</sup> siècle (Bérard 1755 ; Blanchet 1756) et la technique des *sons à caractère* (c'est-à-dire des consonnes doublées) que nous avons décrite ailleurs<sup>7</sup>, témoigne de la grande importance des sons consonantiques. En revanche, les voyelles qui restent intégrées dans la chaîne sonore, sans être trop affectées par une emphase dynamique, ne sont guère discutées chez les mêmes auteurs. Le traitement du souffle dépend de la quantité et de l'intensité souhaitée des sons qui deviennent alors « nourris », « moelleux », « majestueux », « tendres », « légers et gracieux » (Blanchet 1756 : 38)<sup>8</sup>.

Tout autre est le chant italien. Les voyelles sont fleuries, leur flexibilité dynamique porte l'émotion, soutenue par la quantité vocalique. C'est ainsi qu'André Maugars peut constater en 1639 que les chanteurs italiens ont des flexions de voix inconnues aux Français. Le chant italien est extraverti et aborde directement l'auditeur.

---

6. « *Più armoniosa e più varia, capace di atteggiarsi a seconda dei movimenti dell'animo, musicale e pittoresca, sarebbe meno sorda a rispondere all'ingegno de' Francesi, e suonerebbe più grata all'orecchio de' forestieri* » (Algarotti 1750 : 12).

7. Schweitzer (2018) ; Pillot-Loiseau & Schweitzer (2021).

8. Une caractérisation typique du chant français par les Italiens, formulée dans le « Ballet de la raillerie » de Jean-Baptiste Lully (1659 : 13) est alors : « Toy par tes nottes languissantes. Tu pleures plus que tu ne chantes ».

L'idéal est une voix *rein und hell* (« pure et claire », Tosi & Agricola 1757 : 21). Le chanteur prête toute son attention à la sonorisation des sons vocaliques<sup>9</sup>, au point que Jean-Louis Le Cerf de Viéville (1705 (II) : 74) critique que « les Italiens, en chantant, ne prononcent rien »<sup>10</sup>. Effectivement, Tosi & Agricola (1757 : 42) insistent sur la juste prononciation des voyelles, mais ils ne donnent aucun conseil concret pour la sonorisation des consonnes.

Par le traitement du matériau sonore du texte (voyelles et consonnes), chaque composition baroque rend (idéalement) compte des particularités de la langue : soutenue par une technique adaptée, le chant semble *naturel*. Chaque langue peut alors avoir un chant qui lui correspond naturellement. Mais quel chant est le plus naturel ? Pour Rousseau, la réponse est simple. Elle ressort de la distinction phonético-anatomique des langues qu'il détaille dans son *Essai sur l'origine des langues* (1755) et dans son texte manuscrit intitulé *L'origine de la mélodie* :

La mélodie, en imitant les inflexions de la voix, exprime les plaintes, les cris de douleur ou de joie, les menaces, les gémissemens ; tous les signes vocaux des passions sont de son ressort. Elle imite les accens des langues, et les tours affectés dans chaque idiôme à certains mouvemens de l'ame : elle n'imité pas seulement, elle parle ; et son langage inarticulé, mais vif, ardent, passionné a cent fois plus d'énergie que la parole même. (Rousseau 1817 [1781] : 51)

Suivant les raisonnements de Rousseau, on peut dire qu'une chaîne sonore, constituée par un tissu vocalique, se caractérise par

---

9. Chez l'auteur allemand Quantz (1752 : 323), on trouve une confirmation (sans jugement de valeur) qui est en même temps une explication pour ce fait. Selon lui, le style français est « *mehr sprechend als singend* » (« plus parlé que chanté »). Or, ce sont les consonnes qui assurent – toujours – la bonne compréhension du texte, et cela d'autant plus pour un auditeur qui a l'habitude que ces sons soient mis en évidence.

10. Les auteurs insistent sur l'importance de la bonne compréhension du texte et indiquent de manière générale que dans des grandes salles ou à l'extérieur, les consonnes doivent être prononcées plus fort que lorsque l'on parle ordinairement (« *Die Mitlauter insbesondere müssen [...] in großen Gebäuden oder offenen Orten, schärfer als in gemeiner Rede, ja fast übermäßig scharf ausgesprochen werden ; doch mit genauer Beobachtung der ihnen eigenen Härte oder Weichheit.* » Tosi et Agricola 1757 : 138).

sa continuité. Elle est naturelle et naturellement expressive. Les consonnes (articulations) en revanche interrompent ce tissu continu. Sans ajouter à l'expressivité naturelle des expressions vocaliques, elles sont un supplément purement mécanique, artificiel et intellectuel. La chaîne sonore devient discontinue. Plus une langue – et par conséquent son chant – s'appuie sur les sons consonantiques, plus elle s'éloigne de l'expression naturelle. Plus elle met en avant les voyelles, plus elle est naturellement expressive. Dans la vision des rousseauistes, l'adjectif *naturel* désigne une qualité communicative entre les hommes, fondée sur une énergie primaire<sup>11</sup>.

### 3. Une langue facile et commode : une question pratique

Comme le montre l'exemple de Nathalie Dessay, citée au début, la facilité de l'expression vocalique semble pour les chanteurs aller de pair avec une commodité technique<sup>12</sup>. En effet, pour que la voix chantée sorte librement de l'appareil phonatoire, le canal doit être le plus ouvert possible. Au XVIII<sup>e</sup> siècle, Charles de Brosses (1765 : 145) déjà le constate : « Le chant pour être pur, ne doit sortir qu'à plein canal, de la bouche ouverte, et non d'aucun autre tuyau ou partie de l'instrument. Il ne veut donc que des voyelles simples et franches. » De Brosses conclut que « aussi les voyelles pures, dont la langue italienne est remplie, ont-elles assuré la prééminence à la musique de cette nation » (Brosses 1765 : 146)<sup>13</sup>. Aussi Tosi & Agricola (1757 : 30) déclarent-ils catégoriquement : « *diejenige Sprache [ist] zum Singen die beste, welche die wenigsten den freyen Ausgang der Töne hindernden Wörter hat* » (« la meilleure langue pour le chant est celle qui contient le moins de mots empêchant la libre sortie des sons »).

Par les moyens de l'instrumentation, les études modernes ont pu montrer que la couleur de la voix est profondément influencée par l'ouverture du canal et que la portée de la voix est due au formant du

---

11. Cf. Lambert 2005, Kintzler 2006.

12. Le *Petit Robert* donne comme une explication du terme « simple » que la chose désignée est « aisée à utiliser », « commode ».

13. Cf. Schweitzer 2019.

chanteur, une caractéristique attestée chez les comédiens et chez les chanteurs lyriques professionnels sur toutes les voyelles. Ce formant correspond à un pic de résonance, située dans la meilleure zone de sensibilité de l'oreille (vers 3000 Hz, Sundberg 1987). Par l'activation de cette résonance vocalique, le chanteur se produit sans effort (Pillot-Loiseau & Vaissière 2009), il passe facilement au-dessus d'un orchestre et garde la possibilité d'intensifier et d'accentuer (par les sons vocaliques) son chant comme il le souhaite<sup>14</sup>.

Cette technique lui permet non seulement de briller, mais aussi de nuancer la couleur de sa voix et de jouer sur toute la palette des sons expressifs dont il dispose. Le prix en est (presque toujours) la sous-articulation des consonnes et par conséquent une compréhension difficile du texte chanté (Scotto di Carlo 1994, Révis 2013). L'expressivité s'exprime par le son même, par sa couleur, par l'accent et par l'intensité de la voix, c'est-à-dire pour l'auditeur par une compréhension intuitive du sens global, et non du sens spécifique et nuancé du texte. Vu sous cet angle, le chant lyrique s'approche de la musique instrumentale ou absolue qui, dépourvue de mots, ne peut être expressive que par les sons musicaux. On parle ici de ce que Leipp & Castellengo (1969) appellent la « super-forme » globale du son, une forme vocale qui transmet son sens général, même si quelques facteurs comme le tempo et l'intonation ou, ici, la relation de la longueur des sons vocaliques et consonantiques entre eux, varient considérablement<sup>15</sup>. Cette « super-forme » globale est d'autant plus facilement acceptée que, comme le note Malmberg (1971 : 337) pour les Italiens, « le sentiment linguistique de la distinction d'ouverture (de 'saturation') n'est pas très fort chez les locuteurs natifs », fait compréhensible au regard de la grande variété dialectale avec différents systèmes phoniques, existant aujourd'hui encore en Italie.

Dès 1776, les Français peuvent lire la méthode de chant, écrite par l'Italien Giambattista Mancini, dans une traduction française et ils y apprennent que le chanteur arrive à bien développer la voix et sa portée en travaillant notamment les voyelles *a* et *e* « parce qu'elles sont les plus sonores » (1776 : 28). L'école italienne de chant fait

---

14. Cf. Castellengo (2015 : 209).

15. Cf. Castellengo (2015 : 154-156).

son chemin en France, et en 1796, Johannes Paul Aegidius Martini déclare que le grand nombre de voyelles dans la langue italienne fait qu'elle est « tres sonore par elle-même et conséquemment très favorable à la musique » (Martini 1796 : préface). En 1798, Florido Tomeoni revendique ouvertement d'adopter la technique italienne pour la manière de travailler la voix chantée : « L'école française surtout apprend à porter la voix de cette manière ; il est à désirer qu'en se déterminant à fraterniser avec l'école italienne, elle en adopte aussi les principes » (Tomeoni 1798 : 19-20). Malgré un changement de l'écriture musicale pour le chant qui s'effectue au cours de la deuxième moitié du XVIII<sup>e</sup> siècle et qui démontre d'une nouvelle écoute de la langue française (par exemple chez André-Ernest-Modeste Grétry, 1741-1813), on assiste à l'adoption de plus en plus générale de la technique italienne du chant pour les compositions dans n'importe quelle langue. Cette tendance se renforce encore au cours du XIX<sup>e</sup> siècle. Les chanteurs profitent de la possibilité de s'appuyer largement sur les sons vocaliques et aujourd'hui, les études montrent une augmentation considérable de la durée moyenne des voyelles chez les chanteurs lyriques (Scotto di Carlo & Autesserre 1992, Scotto di Carlo 1993<sup>16</sup>).

Cette évolution accompagne un changement épistémologique plus large qui concerne une profonde transformation de la conception de la musique. De la servante de la parole (le fameux « *prima le parole, dopo la musica* »), elle s'élève au rang d'une langue même :

C'est un langage plus touchant, plus énergique que le langage ordinaire ; c'est une langue qui n'est composée que de simples sons, au lieu de renfermer des sons articulés et joints ensemble pour former des mots. Par là il est susceptible de bien plus d'expression que le langage ordinaire, puisqu'il peut offrir bien plus de variété. (Lacépède 1785 : 24)<sup>17</sup>

---

16. Le texte de Martienssen-Lohmann (2010 : 115), s'adressant aux chanteurs et paru en 2010, revendique par exemple très clairement cette technique.

17. Cf. Louis de Jaucourt dans l'article « opéra » dans *l'Encyclopédie* (1765) : « L'intelligence des sons est tellement universelle, qu'elle nous affecte de différentes passions, qu'ils représentent aussi fortement, que s'ils étoient exprimés dans notre langue maternelle. Le langage humain varie suivant les diverses nations. »

Pour être compréhensible, la musique n'a donc besoin que des sons simples, des voyelles, comme les prête en grand nombre la langue italienne.

On peut ajouter un deuxième aspect de simplicité : pour les chanteurs étrangers, l'italien reste une langue plus facilement réalisable, car l'orthographe et la prononciation s'y ressemblent fortement. Pour le français au contraire, l'Abbé Vallart (1744 : 2) constate l'existence de « beaucoup plus de sons simples que [...] de caractères simples », et « assez souvent le même son [est exprimé] par des caractères différents », tandis que « des caractères précisément les mêmes [peuvent exprimer] des sons entièrement différents ».

Jean Écalle et Annie Magnan (2002 : 120) nous apprennent que « les points caractéristiques des différences entre les langues relèvent du degré de transparence de l'orthographe, du type de la simplicité/complexité et l'ouverture/fermeture des structures syllabes ». Ainsi, déjà les enfants italiens maîtrisent beaucoup plus tôt la correspondance entre graphie et prononciation de leur langue que ceux dotés d'un système d'écriture plus complexe (car moins transparent au niveau de cette correspondance comme c'est le cas pour le français). De la même manière, pour les locuteurs étrangers, moins familiers avec le système phonétique de la langue, les erreurs sont largement réduites dans la lecture d'un texte en italien. Katz et Frost (1992) parlent de « profondeur de l'orthographe » pour désigner le degré de complexité de la nature entre graphèmes et phonèmes d'une langue et ils rangent les langues dans une progression allant du simple au complexe selon l'ordre suivant : italien, allemand, français, anglais, hébreu.

L'italien est alors *facile* sous deux angles : d'abord au niveau de la réalisation du texte écrit, et ensuite, mécaniquement, par la facilité de le faire sonner et de le rendre expressif, grâce à sa richesse en voyelles. Chanter en italien aide alors à construire l'image émotionnelle, soit du chanteur lui-même, soit du personnage qu'il interprète : il s'agit d'un *ethos* discursif qui a pour objectif d'émouvoir autrui<sup>18</sup>.

---

18. Cette perspective de l'ethos a été développé par Amossy (1999).

#### 4. L'évidence : une question idéologique

Plus généralement, cet ethos est théorisé à partir du xvii<sup>e</sup> siècle sous la notion du « génie de la langue » (Hassler 2012). Il s'agit d'un débat sur les particularités des langues, leur énergie, leur harmonie, leur clarté d'expression, leur simplicité ou leur richesse. Le génie de la langue est un « idole, comme on le croit communément superbe, intraitable, autosuffisant, dédaigneux de toute communication ou commerce » (Cesarotti 1785/1800 : 101)<sup>19</sup>. Il concerne le lexique et la syntaxe aussi bien que les paramètres prosodiques et la rhétorique typique pour chaque langue :

Que ce génie est biforme, et peut être divisé en deux parties, dont l'une peut être appelé un génie grammatical, et l'autre rhétorique : le premier dépend de la structure mécanique des éléments de la langue et de leur syntaxe ; et l'autre du système général des idées et des sentiments, prédominant dans les différentes nations. (Cesarotti 1785/1800 : 101-102)<sup>20</sup>

Comme ici chez Cesarotti, le génie de la langue est régulièrement mis en relation avec le caractère de ses locuteurs. François Charpentier (1683 : 89) confirme qu'il est « naturel de mesurer le génie des peuples par les qualitez du langage dont ils se servent ». Dans ce sens, D'Alembert oppose dans *La Liberté de la musique* (1759, chap. XVIII) « la chaleur et la variété des Italiens » à « notre monotonie », « notre froideur et [...] notre indigence ». Les caractérisations des Italiens comme « extaovertis, affich[ant] leur joie et leur douleur, sentimentaux, aimant l'informel », et des Français comme « maîtres de soi, froids, d'une perspicacité lucide, amis de la forme » (Harnoncourt 1984 : 193) se retrouvent dans les

---

19. « *Dalla riunione d'ambidue queste parti formasi ciò che si chiama il genio delle lingue ; idolo, come si crede comunemente, superbo, intrattabile, sufficiente a sè stesso, sdegnatore di qualunque comunicazione o commercio* », Cesarotti (1785/1800 : 101). Je remercie Norma Romanelli de son aide pour la traduction du texte italien.

20. « *Che questo genio è biforme, e può distinguersi in due, l'uno de'quali può chiamarsi genio grammaticale, e l'altro rettorico: il primo dipende dalla struttura meccanica degli elementi della lingua, e dalla loro sintassi; l'altro dal sistema generale dell'idee e dei sentimenti che predomina nelle diverse nazioni* », Cesarotti (1785/1800 : 101-102).

descriptions des langues, et cela, dans les textes linguistiques comme dans celles sur la musique des deux peuples.

La langue française se désigne par « le Génie de l'exactitude & les regles de l'Elegance » (Charpentier 1683 : 89). La pensée y est exprimée avec précision et clarté. Non seulement elle rejette l'emploi excessif des superlatifs qui sont fréquents dans les langues italienne et espagnole (Bouhours 1675 : 293), mais sa chaîne sonore est aussi régulière et équilibrée et, par conséquent, ne détourne pas l'attention de l'auditeur du sens exprimé par la construction des mots<sup>21</sup>. L'utilisation fréquente de la voyelle *e*, « beaucoup plus douce que l'*a* » (Corneille 1704 : 162), contribue à la création d'une sonorité douce et chatoyante à l'oreille :

On observe une différence marquée & constante entre le génie ou le goût d'une Nation & la Langue qu'elle parle : comme chaque voyelle a un son particulier, plus fort ou plus foible, chaque Nation, selon son inclinaison dominante, affecte de se servir des voyelles qui conviennent le plus à son humeur.

Les Italiens font [...] un grand usage des terminaisons en *A* & en *O*, parce qu'ils aiment naturellement l'amplification & l'exagération. Les Français, qui aiment ce qui est naturel<sup>22</sup>, & fuient l'affectation, se servent volontiers de la lettre *E*, dont la prononciation est plus douce. (Le Pilleur d'Apligny 1779 : 55)

Comme la langue française, la musique française est à l'âge baroque dictée par des règles précises et peu de place est laissée à l'improvisation. L'objectif est le raffinement, accompagné d'une transparence absolue au niveau de la forme. Harnoncourt (1984 : 202) parle d'un « plaisir hautement recherché, destiné à des auditeurs fins, sensibles à des productions artistiques raffinées et spirituelles ». On pourrait aussi dire qu'il s'agit d'une musique complexe et pas toujours facile d'accès. A la fin du xviii<sup>e</sup> siècle, Antoine de Rivarol déclarera même dans son *Discours sur l'universalité de la langue française* que :

Ce n'est point, comme on l'a dit, parce que les mots français ne sont pas sonores que la musique les repousse, c'est parce qu'ils offrent l'ordre et la

---

21. Cf. Meschonnic (1997 : 210-211).

22. On note en passant une autre signification attribuée à l'adjectif *naturel* chez l'auteur.

suite, quand le chant demande le désordre et l'abandon. La musique doit bercer l'âme dans le vague et ne lui présenter que des motifs. (Rivarol 1991 [1784])

Autrement dit : selon la bipartition de Cesarotti (*supra*), dans le génie de la langue française, la syntaxe prime sur la rhétorique. Toutefois, dans le même texte, Rivarol critique aussi la sonorité de la prose italienne, dévalorisant en quelque sorte son premier argument. Selon lui, en italien,

Toutes les lettres se prononcent, & roulant toujours sur des sons pleins, [l'italien] se traîne avec trop de lenteur : son éclat est monotone, l'oreille se lasse de sa douceur, & la langue de sa mollesse ; ce qui peut venir de ce que chaque mot étant harmonieux en particulier, l'harmonie de tout ne vaut rien. (Rivarol 1991 [1784])

On peut résumer cet argument comme suit : comme le français, l'italien est une langue harmonieuse et douce, mais ces caractéristiques s'appuient en italien sur les mots isolés, et en français sur l'enchaînement des mots. Si la prononciation du français est « moins éclatante » que celle de l'italien, elle reçoit notamment par le son du *e* muet « une harmonie légère qui n'est qu'à elle » (Rivarol 1991 [1784]). Ces constats retrouvent nos considérations sur le *naturel* des langues (*supra*).

Rousseau, contrairement à Rivarol défenseur de la langue italienne, attribue quatre caractéristiques à cette langue. Premièrement, elle est douce, car elle n'utilise que peu de consonnes et notamment de combinaisons de consonnes. Deuxièmement, elle est sonore, et cela, grâce à ses voyelles « éclatantes ». Elle est aussi harmonieuse par sa quantité variée et par sa prosodie. Enfin, elle est « accentuée plus qu'aucune autre », c'est-à-dire expressive par ses sons vocaliques (*supra*). Or, ce sont exactement ces qualités qui sont « les plus convenables au chant » (Rousseau 1753) : comme la prosodie de la langue italienne est très variée, elle est aussi à l'origine de mélodies intéressantes et riches. Son adéquation pour le chant est tout simplement une évidence<sup>23</sup>.

---

23. Le *Petit Robert* donne (entre autres) comme explication du terme « simple » que la chose désignée « présente une évidence ».

Par son argumentation en faveur de la musique italienne, Rousseau avance en quelque sorte les quatre critères d'une langue « parfaite », évoqués par Daniel Jenisch lors du Prix de l'Académie des sciences de Berlin en 1792 (Jenisch 1792 : 5) : premièrement son *Reichtum* (sa richesse), c'est-à-dire la variété de sa prosodie ; deuxièmement sa *Regelmäßigkeit* (son analogie) qui est à l'origine de sa douceur ; troisièmement sa *Kraft* (sa force), c'est-à-dire sa sonorité pleine ; et quatrièmement son *Harmonie* (son harmonie) qui fournit toute son expressivité aux voyelles. Le tableau 1 compare les caractéristiques mises en avant par Rousseau pour la langue italienne, avec celles indiquées par Jenisch pour une langue parfaite.

Rousseau (l'italien)	Jenisch (la langue parfaite)
Harmonieux (une quantité et une prosodie variées)	<i>Reichtum</i> (richesse)
Doux (peu de consonnes)	<i>Regelmäßigkeit</i> (analogie, régularité)
Sonore (des voyelles éclatantes, une sonorité pleine)	<i>Kraft</i> (force)
Accentué (l'expressivité des voyelles)	<i>Harmonie</i> (harmonie)

Tableau 1 : Les caractéristiques de la langue italienne (Rousseau 1753) et de la langue parfaite (Jenisch 1792)

## 5. Conclusion ou : Une langue simple pour le chant ?

En vérité, ce qui distingue les langues, l'italien et le français, ce ne sont pas leurs capacités expressives, mais c'est leur manière particulière d'être expressive. Chaque argumentation en faveur d'une langue ou d'une autre, même si elle s'appuie sur des observations comme la fréquence ou la qualité des différents sons des langues, repose sur des considérations esthétiques, et donc subjectives. Le chant dépend directement des particularités phoniques d'une langue. Il puise dans ce que la langue utilisée lui offre et il met ainsi en avant des caractéristiques sonores et prosodiques précises.

Posons alors, pour finir, la question que pourrait être, sous cet angle, la définition de la simplicité d'une langue par rapport au chant

et comment on peut expliquer l'image de l'italien comme la *meilleure* langue pour le chant.

Il semble que la préférence de l'italien pour le chant est fondée sur deux types d'arguments, l'un est technique (la facilité de faire sonner et de porter la voix), et l'autre esthétique (le caractère et l'énergie des sons et de leurs enchaînements ainsi que l'*ethos* discursif véhiculé par la langue). La naissance de la suprématie de l'italien est intimement liée à un contexte musical et esthétique, et elle est accompagnée d'une large discussion autour de la question de la bonne langue pour le chant en particulier, et de la bonne musique en général. Pour comprendre l'évolution épistémologique de la fin du XVII<sup>e</sup> siècle jusqu'à aujourd'hui, un regard dans l'histoire de l'esthétique musicale s'impose.

Nous avons vu qu'à l'âge baroque et classique (1600-1800), la question de savoir quelle est la meilleure langue pour chanter n'est, au fond, pas décidée. Charles Burney constate d'ailleurs ce fait de manière très pragmatique :

Si l'on considère que la musique française est bonne, et son expression naturelle et agréable, la musique italienne doit alors être mauvaise ; ou inversement, si la musique italienne a tout ce qu'une oreille non pourrie gâtée, bien exercée peut souhaiter, on peut alors présumer que la musique française ne donnera pas autant de plaisir à une telle oreille. (Burney 1772 : 12-13)<sup>24</sup>

Le choix de l'italien ou du français comme langue chantée est lié à l'esthétique d'une certaine culture musicale, un fait qui se justifie largement par le lien étroit qu'entretiennent langue et musique à l'époque.

Tout change avec la nouvelle esthétique romantique. Une école universelle s'installe en Europe occidentale, celle qui poursuit la technique du *belcanto* italien. Une première tendance à l'italianisation de l'opéra en France s'annonce avec Jean-Philippe Rameau (Cy 1980), et elle se manifeste pleinement dans les œuvres de Christophe

---

24. « *Ist die französische Musik gut, und ihr Ausdruck natürlich und gefällig, so muß die italiänische schlecht seyn ; oder umgekehrt, wenn die italiänische Musik alles hat, um ein unverwöhntes, wohl geübtes Ohr wünschen kann : so läßt sich nicht vermuthen, daß die französische Musik, einem solchen Ohre eben so viel Vergnügen machen werde* », Burney (1772 : 12-13).

Willibald Gluck. De l'image de la sonorité et de l'expressivité d'une langue et d'une sensation des conditions de la portée de la voix chantée, on passe à une technique concrète. En France, la publication d'une collection de solfèges de compositeurs italiens (Bêche & Lévesque 1772) marque un moment décisif dans cette évolution. L'enseignement du chant dans l'École Royale de Chant, fondée en 1784 par François-Joseph Gossec, repose également sur la méthode italienne. L'italianisation du chant français culmine dans la publication de la *Méthode de chant du Conservatoire de musique contenant les principes du chant des exercices pour la voix, des solfèges tirés des meilleurs ouvrages anciens et modernes* en 1802/1803, un ouvrage de méthodologie italienne, traduit dans de nombreuses langues. Il pose encore les fondements de la méthode de chant la plus influente du XIX<sup>e</sup> siècle, le *Traité complet de l'art du chant* de Manuel Garcia (Paris 1840/1847).

Avec Garcia (1855), le premier à avoir observé en 1855 le comportement des cordes vocales d'un chanteur lors des changements de registre, on passe à une technique scientifiquement fondée, dont les bases sont valables jusqu'à aujourd'hui. Au centre de cette technique italienne ou, dorénavant, universelle, se trouve la perfection de la portée, de la brillance et des nuances de la voix. L'orchestre prenant plus de poids dans la composition musicale, cette formation intense de la voix devient nécessaire, voire indispensable. Pour s'entraîner, on reprend la méthode du *solfège*, basée sur l'exercice des mélismes vocaliques, et destinée à former et à fortifier la voix.

S'y ajoutent deux évolutions dans le monde musical. Premièrement, les instruments deviennent plus forts et plus puissants, et deuxièmement, les instrumentistes dans les orchestres plus nombreux. A la fin du XVIII<sup>e</sup> siècle, Florido Tomeoni (1798 : 38) s'apitoie sur « le pauvre chanteur » qui « est obligé de s'égosiller, non pour se faire comprendre du public, mais simplement pour se faire entendre ». On comprend que le volume de la voix devient un des critères principaux pour la réussite d'un chanteur. Tout cela favorise non seulement la technique, mais aussi le choix de la langue italienne. Comme langue parlée par une nation « extrovertie » qui montre facilement ses émotions (*supra*), ses fortes capacités expressives dans le chant semblent évidentes. L'image du caractère

d'une langue et du peuple qui la parle se manifeste dorénavant dans le fait que l'italien est considéré non seulement comme une langue musicale, mais comme « la langue reine dans le monde de la musique et de l'Opéra » (*supra*). Riche en voyelles, elle permet de développer la sonorité vocale. Besançon van Oyen (1851 : 26-27) recommande de faire les premiers exercices sur les voyelles *a* et *o*, ce sont elles qui « offrent le plus d'avantages » au chanteur : « *O* d'abord, parce qu'il agrandit l'arrière-bouche et développe, par cela même la voix ; *A* ensuite, parce qu'il empêche le son d'être sourd, qu'il le rend plus extérieur et lui donne plus d'éclat ».

Dans cette entreprise, la langue italienne peut être considérée comme plus simple que les autres, car elle permet au chanteur de se concentrer sur le son vocalique. Dans le but de déployer sa voix, le chant italien assure directement la prégnance à la voix, tandis que d'autres langues, plus riches en consonnes, demandent des manipulations plus complexes de l'appareil phonatoire et un autre type de maîtrise de la voix, considéré comme moins simple d'accès.

Sous cet angle, une langue simple pour le chant peut être définie comme une langue qui facilite et qui favorise l'expression du chanteur par sa sonorité intrinsèque.

## Bibliographie

### Sources primaires

- Anonyme. 1802/1803. *Méthode de chant du Conservatoire de musique contenant les principes du chant des exercices pour la voix, des solfèges tirés des meilleurs ouvrages anciens et modernes*. Paris.
- Alembert, Jean le Rond d'. 1759. *De la liberté de la musique*. Amsterdam : Chatelain.
- Algarotti, Francesco. 1750. *Saggio sopra la lingua francese*. s. l.
- Bèche, Jean-Louis & Jean-Louis Pierre Lévesque. 1772. *Solfèges d'Italie avec la basse chiffrée*. Paris.
- Bouhours, Dominique. 1675. *Remarques nouvelles sur la langue française*. Paris : Marbre.
- Brosses, Charles de. 1765. *Traité de la formation mécanique des langues et des principes physiques de l'étymologie*. Paris : Vincent & Dessaint.

- Burney, Charles. 1772. *Tagebuch einer Musikalischen Reise durch Frankreich und Italien*. Hambourg : Bode.
- Cesarotti, Melchiorre. 1785/1800. *Saggio sulla filosofia delle lingue applicato alla lingua italiana*. Pisa : Dalla Tipografia della società lett.
- Charpentier, François. 1683. *De l'excellence de la langue française*. Paris : Bilaine.
- Corneille, Thomas. s. d. *Observations de l'Académie française sur les remarques de M. de Vaugelas*. 2<sup>e</sup> éd., La Haye : Honoré & Johnson.
- Garcia, Manuel. 1847. *Traité complet de l'art du chant*. Paris.
- Jaucourt, Louis de & Jean-Jacques Rousseau. 1765. Chant. *Encyclopédie*, éd. par Denis Diderot & Jean Le Rond d'Alembert. Paris : Briasson, David, Le Breton & Durand. 494-496.
- Jenisch, Daniel. 1796. *Philosophisch-kritische Vergleichung und Würdigung von vierzehn älteren und neueren Sprachen*. Berlin : Maurer.
- Le Cerf de La Viéville, Jean-Louis. 1705. *Comparaison de la musique italienne et de la musique française*. Bruxelles : Foppens.
- Le Pileur d'Apligny. 1779. *Traité sur la musique et sur les moyens d'en perfectionner l'expression*. Paris : Demonville & Saugrain.
- Lully, Jean-Baptiste. 1659. *Ballet de la raillerie*, Livret. Paris : Ballard.
- Mancini, Giambattista. 1776. *L'Art du chant figuré, traduit de l'italien par M. A. Desaugiers*, Vienne. Paris : Cailleau, Durand, Lacombe & Duchesne.
- Martini, Johann Paul Aegidius. 1796. *Melopée moderne ou L'Art du chant réduit en principes*. Paris : Nadermann.
- Maugars, André. 1638. *Response faite à un curieux sur le sentiment de la musique d'Italie*. Rome.
- Quantz, Johann Joachim. 1752. *Versuch einer Anleitung die Flöte traversière zu spielen*. Berlin : Voss.
- Rivarol, Antoine de. 1991 [1784]. *De l'universalité de la langue française*. Paris : Distique.
- Rousseau, Jean-Jacques. 1753. *Lettre sur la musique française*. s. l.
- Rousseau, Jean-Jacques. 1995 [1755]. *L'origine de la mélodie*. Texte établi et édité par Bernard Gagnebin & Marcel Raymond. *Jean-Jacques Rousseau, Œuvres complètes*, t. 5 (*Écrits sur la musique, la Langue et le Théâtre*). Paris : Gallimard (Bibliothèque de la Pléiade). 331-346.
- Rousseau, Jean-Jacques. 1817 [1781]. *Essai sur l'origine des langues, où il est parlé de la Mélodie, et de l'Imitation musicale*. Édition électronique établie par Daniel Banda à partir de l'édition des Œuvres (t. 4). Paris : A. Belin. 501-543. [[http://classiques.uqac.ca/classiques/Rousseau\\_jj/essai\\_origine\\_des\\_langues/origine\\_des\\_langues.pdf](http://classiques.uqac.ca/classiques/Rousseau_jj/essai_origine_des_langues/origine_des_langues.pdf), consulté le 13/01/2023].
- Sévigny, Mme de. 1806. *Lettres de Madame de Sévigné à sa fille et à ses amis*, t. 2. Edition de Ph. A. Grouvelle. Paris : Bossange.

- Tomeoni, Florido. 1799. *Théorie de la musique vocale*. Paris.
- Tosi, Pier Francesco & Johann Friedrich Agricola. 1757. *Anleitung zur Singkunst*. Berlin : Winter.
- Vallart. 1744. *Grammaire française*. Paris : Dessaint & Saillant.
- Walter, Johann Gottlieb. 1732. *Musikalisches Lexikon oder musikalische Bibliothek*. Leipzig : Deer.

## Sources secondaires

- Amossy, Ruth. 1999. *Images de soi dans le discours. La construction d'un ethos*. Lausanne : Delachaux & Niestlé.
- Canepari, Luciano. 1976. The dialect of Venice. *Journal of the International Phonetic Association*. 67-76.
- Cannone, Belinda. 2006. Littérature et musique. *Histoire de la France littéraire. Classicismes XVII<sup>e</sup> – XVIII<sup>e</sup> siècles*, dir. Par J.-Ch. Darmon & M. Delon. Paris : PUF. 518-546.
- Castellengo, Michèle. 2015. *Ecoute musicale et acoustique*. Paris : Eyrolles.
- Cyr, Mary. 1980. Eighteenth-Century French and Italian Singing : Rameau's Writing for the Voice. *Music & Letters* 61 (3-4). 318-337.
- Delattre, Pierre. 1953. Les modes phonétiques du français. *French Review* 27(1). 59-63.
- Écalle, Jean & Annie Magnan. 2002. *L'apprentissage de la lecture : Fonctionnement et développements cognitifs*. Paris : Colin.
- Garde, Paul. 1968. *L'accent*. Paris : PUF.
- Harnoncourt, Nikolaus. 1984. *Le discours musical. Pour une nouvelle conception de la musique*. Texte traduit de l'allemand par Dennis Collins. Paris : Gallimard.
- Hassler, Gerda. 2012. La description de génie de la langue dans les grammaires françaises et les grammaires d'autres langues. *Vers une histoire générale de la grammaire française. Matériaux et perspectives*, dir. par Bernard Colombat, Jean-Marie Fournier & Valérie Raby. Paris : Champion. 193-209.
- Katz, Leonard & Ram Frost. 1992. The reading process is different for different orthographies : The orthographe depth hypotheses. *Orthography, Phonology, Morphology and Meaning*, dir. par Ram Frost & Leonard Katz. Amsterdam : Elsevier Science Publishers. 67-84.
- Kintzler, Catherine. 2006. *Poétique de l'opéra français — de Corneille à Rousseau*. Paris : Minerve.
- Lambert, Alain. 2005. *Musique, philosophie et littérature chez Rousseau (suivi d'un entretien sur la musique avec le musicien)*. [[https://www.musicologie.org/publiem/lambert\\_rousseau.html](https://www.musicologie.org/publiem/lambert_rousseau.html)], consulté le 19/02/2023].

- Leipp, Emile & Michèle Castellengo. 1969. L'intelligibilité de la parole dans le chant. *L'Oreille, juge de la qualité sonore*. Paris : Chiron. 15-34.
- Malmberg, Bertil. 1971. *Phonétique générale et romane. Études en allemand, anglais, espagnol et français*. La Hague & Paris : Mouton De Gruyter.
- Martienssen-Lohmann, Franziska. 2010. *Der wissende Sänger*. Zurich : Atlantis.
- Meschonnic, Henri. 1997. *De la langue française. Essai sur une clarté obscure*. Paris : Hachette.
- Pierret, Jean-Marie. 1994. *Phonétique historique du français et notions de phonétique générale*. Louvain-La-Neuve : Peeters.
- Pillot-Loiseau, Claire & Claudia Schweitzer. 2021. Le doublement de consonnes dans la musique baroque française : objectifs, technique et phonétique. *Histoire de la description de la parole : de l'introspection à l'instrumentation*, dir. par Christelle Dodane & Claudia Schweitzer. Paris : Champion. 319-342.
- Pillot-Loiseau, Claire & Jacqueline Vaissière. 2009. La portée de la voix parlée et chantée : aspects scientifiques et rééducatifs. *Unadreo, La Voix dans tous ses maux*. Ortho Edition. 243-249.
- Ramus, Franck. 1999. *Rythme des langues et acquisition du langage*. Thèse de doctorat. EHESS.
- Robert, Paul. 2010. *Le Nouveau Petit Robert 2010*. Paris : Sejer.
- Roegiest, Eugeen. 2009. *Vers les sources des langues romanes. Un itinéraire linguistique à travers les Romania*. Leuven & Le Hague : Acco.
- Révis, Joana. 2013. *La voix et soi. Ce que notre voix dit de nous*. Paris : De Boeck.
- Rogers, Derek & Luciana d'Arcangeli. 2004. L'Italien. *Journal of the International Phonetic Association*. 117-121.
- Saffi, Sophie. 2001. Syntaxe et prosodie en italien et en anglais. *Italies* 5. 211-234.
- Sandford, Sally A. 2005 [1995]. A Comparison of French and Italian Singing in the Seventeenth Century. *Journal of Seventeenth-Century Music* 1(1). [<https://sscm-jscm.org/v1/no1/sanford.html>, consulté le 19/03/2023].
- Schweitzer, Claudia. 2018. *Parole et Chant. Histoire des théories du son du français à l'âge classique (xvii<sup>e</sup> et xviii<sup>e</sup> siècles)*, Thèse de doctorat. Université Sorbonne nouvelle - Paris 3.
- Schweitzer, Claudia. 2019. Une bonne langue pour chanter? Réflexions sur les caractéristiques phonétiques des langues et sur le chant baroque. *History and Philosophy of the Language Sciences*. [<https://hiphilangsci.net/2019/11/19/langue-pour-chanter/>, consulté le 19/02/2023].

- Scotto di Carlo, Nicole & Denis Autesserre. 1992. L'organisation temporelle de la syllabe dans la parole et dans le chant. Étude préliminaire. *Travaux interdisciplinaires du Laboratoire Parole et Langage Aix-en-Provence (TIPA)* 14. 189-204.
- Scotto di Carlo, Nicole. 1993. L'auto-contrôle de la voix chez les artistes lyriques. *Médecine des Arts* 3.
- Scotto di Carlo, Nicole. 1994. L'intelligibilité de la voix chantée. *Médecine des Arts* 10.
- Sundberg, Johan. 1987. *The science of the singing voice*. Illinois : Northern Illinois University Press.
- Zedda, Paolo. 2013. Langues chantées : le système allophonique de la diction italienne. *La Revue du Conservatoire*.



# L'émergence du couple terminologique « phrase simple / phrase complexe » dans les grammaires françaises

Sophie Jollin-Bertocchi<sup>1</sup>, Jacques-Philippe Saint-Gerand<sup>2</sup>

<sup>1</sup> CHCSC (EA2448), Université Paris-Saclay, UVSQ

<sup>2</sup> 63380 Miremont

---

## RÉSUMÉ

Le couple simplicité/complexité est appliqué à la phrase pour la langue française comme pour de nombreuses autres langues. Cette distinction – qui ne se confond pas avec la simplicité/complexité syntaxique – émergente au XVII<sup>e</sup> siècle, est apparue en liaison avec la naissance de la notion de subordination, dans le contexte d'un réseau notionnel autour de l'idée de phrase. Elle est ensuite présente dans la pensée de la grammaire scolaire au XIX<sup>e</sup> siècle, mais sans être formalisée en tant que telle, les grammairiens employant d'autres termes (phrase *incomplexe* ou *composée*). Elle recoupe l'opposition entre analyse grammaticale et analyse logique à la faveur de l'opposition entre un enseignement primaire pour tous centré sur la phrase simple, et un enseignement secondaire pour les classes sociales favorisées qui traite de la phrase complexe, rapprochant la grammaire de la rhétorique.

Absente dans les grammaires linguistiques de la première moitié du XX<sup>e</sup> siècle, l'opposition phrase simple / complexe devient, dans la seconde moitié du siècle, explicite dans les grammaires à orientation pratique et didactique, y compris dans les grammaires de référence, dans lesquelles elle joue un rôle structurant – conséquence de l'influence du modèle linguistique génératif. D'un point de vue scientifique, le couple est néanmoins fragilisé par son orientation strictement pédagogique.

La collocation avec le mot *phrase* confère un sens bien spécifique aux deux termes. En théorie, le critère de distinction est clair : un seul verbe et donc une seule proposition pour la phrase simple, plusieurs verbes et par

conséquent plusieurs propositions pour la phrase complexe, laquelle se subdivise alors en deux types, la phrase composée de propositions juxtaposées ou coordonnées et la phrase composée de propositions subordonnées. Or, outre la fonction du verbe, assertive d'un jugement au sens de la grammaire générale, l'on sait que cette typologie pose un problème de nature historiographique dans la mesure où elle met sur le même plan la parataxe et l'hypotaxe, alors que le degré d'implication n'est pas le même dans les deux cas. Au-delà d'une typologie problématique de la structure de la phrase, cette distinction hégémonique est partie prenante dans la difficulté à définir la notion de phrase qui caractérise la linguistique du xx<sup>e</sup> siècle. Elle constitue même l'un des points d'achoppement de cette définition.

### MOTS-CLÉS

grammaire, phrase, proposition, période, rhétorique, logique, syntaxe

### ABSTRACT

The couple simplicity/complexity is applied to the sentence in French grammar, as is the case with many other languages. This distinction - not to be confused with syntactic simplicity/complexity - emerged in the 17th century, in connection with the birth of the notion of subordination, in the context of a notional network centered on the idea of the sentence. It was then present in grammar school manuals in the 19th century, but without being formalized as such, as grammarians used other terms (*incomplete* or *compound sentence*). It overlapped with the opposition between grammatical analysis and logical analysis, and thus justified the opposition between a primary education for all focused on the simple sentence, and a secondary education for privileged social classes that dealt with the complex sentence, bringing grammar closer to rhetoric.

Absent in the linguistic grammars of the first half of the 20th century, the opposition between simple and complex sentences becomes, in the second half of the century, explicit in grammars with a practical and didactic orientation, including reference grammars, in which it plays a structuring role - a consequence of the influence of the generative linguistic model. From a scientific point of view, it is nevertheless weakened by its strictly pedagogical orientation.

The collocation with the word *sentence* confers a very specific meaning to the two terms. In theory, the criterion for distinction is clear: a single verb and therefore a single clause for the simple sentence, several verbs and therefore several clauses for the complex sentence, which is then subdivided into two types, juxtaposed or coordinated clauses and subordinate clauses. However, in addition to the function of the verb as an asserter of a judgement in the sense of general grammar, we know that this typology

poses a historiographical problem insofar as it puts parataxis and hypotaxis on the same level, whereas the degree of implication is not the same in both cases. Beyond a problematic typology of sentence structure, this hegemonic distinction is part of the difficulty in defining the notion of sentence that characterizes 20th-century linguistics. It even constitutes one of the stumbling blocks of this definition.

#### KEYWORDS

grammar, sentence, clause, period, rhetoric, logic, syntax

---

## 1. Introduction

Comme le pose un ouvrage récent, « Il revient à la linguistique, héritière des traditions logiques et grammaticales, d'avoir su fixer des frontières entre chacun de ces trois concepts : proposition, phrase, énoncé » (Neveu 2019 : 5). En-deçà de l'approche linguistique contemporaine, ces trois termes ont déjà fait l'objet de multiples réflexions tout au long de l'histoire des notions grammaticales. Ainsi, associé à la conception de tout énoncé comme expression d'un jugement, le rôle définitoire dévolu au verbe fait-il de la *phrase* un objet longtemps resté grammaticalement problématique, entre proposition logique et période rhétorique expansible. En français comme dans d'autres langues, la collocation des adjectifs *simple* et *complexe* avec le mot *phrase* confère un sens bien spécifique aux deux termes. Depuis le XVII<sup>e</sup> siècle, le critère de distinction est très clair : un verbe, une proposition, une phrase simple, ou plusieurs verbes, plusieurs propositions, une phrase complexe. Cette dernière se subdivisant alors en deux espèces, la phrase à propositions subordonnées (hypotaxe), et la phrase à propositions juxtaposées ou coordonnées (parataxe). La question est alors de savoir comment a pu s'établir à l'époque moderne le double système.

Les catégories *phrase simple* et *phrase complexe* forment un couple conceptuel qui va aujourd'hui de soi, comme en témoignent les ouvrages de grammaire modernes, mais qui reste d'une formalisation néanmoins récente, et qui n'est pas toujours dépourvue d'ambiguïtés épistémologiques. Aujourd'hui, la dichotomie coexiste d'ailleurs avec celle de *simplicité vs complexité syntaxique*. La notion de complexité absolue, objective – par opposition à la complexité

relative, cognitive – est généralement définie en fonction de deux paramètres : le nombre d'éléments et le nombre de connexions entre ces éléments<sup>1</sup>. La complexité syntaxique s'exerce à deux niveaux, interpropositionnel et intrapropositionnel<sup>2</sup>, mais elle concerne aussi la diversité des structures.

On cherchera à retracer ici les conditions d'émergence de ces notions et on se demandera quel est le rendement conceptuel de la terminologie qui s'est imposée au détriment d'autres possibilités.

Nous suivrons un parcours chronologique en commençant par la genèse du XVII<sup>e</sup> au XIX<sup>e</sup> siècle, puis nous montrerons comment s'est faite l'invention d'une tradition avant de terminer par l'état de la critique et les rares solutions proposées.

## 2. Genèse

Les notions de simplicité et de complexité traversent toute l'histoire de la pensée du langage, sans que l'on puisse dégager une réelle formalisation conceptuelle stable. Denys d'Halicarnasse, par exemple, en fait un critère de différenciation de l'atticisme et de l'asianisme dans son *Traité de l'arrangement des mots*. Lorsqu'il critique Isocrate, il souligne que l'essentiel dans le langage de la discussion politique ou du débat public est de rester le plus près possible du naturel, et que la nature réclame que l'expression se mette au service des idées et non l'inverse. De même, lorsqu'il observe Thucydide, il rappelle la nécessité de se garder de ses tournures, ténébreuses, détournées de la construction naturelle. En d'autres termes si, en vertu des lois de la nature, la pensée et la syntaxe possèdent le même principe d'organisation, le discours – pensé ou parlé – n'est correct que s'il adopte une sorte de syntaxe naturelle calquée sur la logique et la dialectique, et donc hostile à tous les procédés d'ornementation du discours, à l'hyperbate comme à l'interpolation :

---

1. Cf. « *the number of parts in a system, of connections between different parts* » (Miestamo 2009 : 81).

2. « Ce qui paraît complexe du point de vue de la syntaxe est dû, la plupart du temps, à une difficulté à extraire d'un énoncé la suite S,V,O [...]. » (Blanche-Benveniste 2013 : 257).

Je croyais que, nous surtout qui nous conformons à la nature, nous devons, dans l'ajustement des mots, respecter ses exigences. (Denys d'Halicarnasse 1978 : 257)

## 2. 1. Simplicité et ordre des mots

Pour ce qui est de la phrase simple, on trouve ainsi, dès l'Antiquité, la problématique de l'ordre des mots dans l'énoncé, balançant entre les impératifs de la raison, de la nature et ceux de l'esthétique ou du goût. Depuis la fin du xvi<sup>e</sup> siècle et tout au long du xvii<sup>e</sup>, l'ordre de nature s'impose donc sous la forme Sujet-Verbe-Objet, qui réunit arguments logiques et critères syntaxiques dans un modèle canonique assurant la spécificité de la langue française contre les diversions du latin. Or ces décennies ne cessent de laisser apparaître quantité de travaux : grammaires, remarques, essais traitant des formes des énoncés comme *locutions*, *périodes*, voire *phrases* sous leurs aspects non théoriques mais pratiques et en quelque sorte fonctionnels (Combettes 2011). Outre Port-Royal, on peut citer Maupas, Oudin, Chiflet, Vairasse d'Allais, Dangeau, mais aussi Bouhours, Ménage, Andry de Boisregard, Vaugelas, La Mothe Le Vayer, etc. Dans ce cadre, le composant rhétorique s'est alors rapidement imposé comme le pivot des discussions relatives aux formes d'expression simples (i.e. *logiques*) ou complexes (i.e. *ornées* et *figurées*) d'une pensée conçue alors comme antérieure à sa formulation et essentiellement immanente. Toute la complexité d'une notion grammaticale en gestation s'est alors nouée dans cette interférence des catégories de la grammaire et de la rhétorique des figures.

## 2. 2. Complexité rhétorique : la période et les figures

Ayant cette capacité à prendre en charge tout ce qui du discours échappe à la grammaire, la rhétorique était à même d'articuler ce passage d'un certain flou conceptuel à une théorisation plus radicale. Grâce aux travaux de Jean-Pierre Seguin et Valérie Raby (cf. Bibliographie), nous disposons désormais à cet égard d'un ensemble d'études approfondies qui permettent de baliser l'émergence progressive du concept grammatical et linguistique

moderne de phrase, à partir des travaux des auteurs précédemment cités, et d'en comprendre les raisons théoriques et pratiques.

La belle simplicité du modèle canonique supposément naturel fut largement discutée, notamment à la fin du XVII<sup>e</sup> et tout au long du XVIII<sup>e</sup> siècle, par les partisans des périodes oratoires dont la force persuasive repose sur le déploiement de stratégies rhétoriques. Le Père de Jouvancy n'écrit-il pas en 1711 :

Voulez-vous, par exemple, faire une période de cette proposition : « Dieu récompense la vertu » ? Exposez les notions impliquées dans les mots DIEU, VERTU, RÉCOMPENSE, et dite : « **Puisque** Dieu est le souverain arbitre de toutes choses, **et** le juge très équitable des bons et des méchants, **on ne peut douter que** les actes de piété et de sagesse **dont** la vertu est la cause ne reçoivent une récompense magnifique, **non seulement** dans cette **vie mais encore** dans l'autre, **qui** est immortelle. (Père Joseph de Jouvancy 1739 : 50)

C'est par la voie des figures rhétoriques que la simplicité syntaxique en vient à s'opposer à la complexité du développement idéologique dans le cadre de la dialectique de la logique des propositions et des formes de leurs énoncés. Une proposition logique simple est capable de donner naissance à une période oratoire complexe. On ne saurait en trouver meilleure illustration dans les pages où Beauzée définit les trois états de la phrase par rapport à la totalité de ses parties :

Par rapport à cette totalité des parties, il y a plénitude, ou défaut, ou réduplication. S'il y a plénitude, c'est l'état naturel de la phrase ; *Scinditur incertum studia in contraria vulgus (m)*. S'il y a défaut ou **Rédondance**, c'est un écart de l'état naturel. L'état où quelque chose manque constitue **l'Ellipse** ; *Longum iter est per præcepta, breve & efficax per exempla (n)* ; la phrase pleine serait, *Iter institutum per præcepta est iter longum, iter institutum per exempla est iter breve & iter efficax*. L'état où il y a plus que la nature n'exige constitue le **Pléonasme** ; *Morte morieris (o)* ; *nequaquam morte moriemini (p)* ; la raison grammaticale semble pouvoir se passer ici du mot *morte*. (Beauzée 1974, t. 2 : 395-96)

## 2. 3. Émergence de la phrase complexe au XVII<sup>e</sup> siècle

La phrase complexe émerge au XVII<sup>e</sup> siècle (Raby 2002, 2017 ; Andrieu 2012) dans la grammaire générale, mais il ne s'agit alors que de la phrase à subordonnée envisagée seulement sous l'angle de

l'argumentation logique. La *Grammaire de Port-Royal*, par exemple, sous la décomposition logique de la « proposition composée », assimile coordination et subordination, tout en noyant relatives et complétives sous la dénomination de « propositions incidentes ». Ainsi dans la citation suivante, comprend-on que l'énoncé *Dieu invisible a créé le monde visible* comporte, deux relatives (incidentes) implicites qui expriment des jugements tacitement formés dans notre esprit par les simples règles de la logique :

Cette union de plusieurs termes dans le sujet et dans l'attribut est quelquefois telle, qu'elle n'empêche pas que la proposition ne soit simple, ne contenant en soi qu'un seul jugement ou affirmation, comme quand je dis : *La valeur d'Achille a été cause de la prise de Troie*. Ce qui arrive toujours toutes les fois que des deux substantifs qui entrent dans le sujet ou l'attribut de la proposition, l'un est régi par l'autre.

Mais d'autres fois aussi, ces sortes de propositions dont le sujet ou l'attribut sont composés de plusieurs termes, enferment, au moins dans notre esprit, plusieurs jugemens, dont on peut faire autant de propositions ; comme quand je dis : *Dieu invisible a créé le monde visible* : il se passe trois jugemens dans mon esprit, renfermés dans cette proposition. Car je juge premièrement que *Dieu est invisible*. 2. *Qu'il a créé le monde*. 3. *Que le monde est visible*. Et de ces trois propositions, la seconde est la principale et l'essentielle de la proposition : mais la première et la troisième ne sont qu'incidentes, et ne font que partie de la principale, dont la première en compose le sujet et la dernière l'attribut. (Arnauld et Lancelot 1803 : 306)

On notera à ce sujet que, si la *Logique ou l'Art de penser* fait référence à deux types distincts de propositions (logiques) incidentes, les *explicatives* et les *déterminatives*, dont les mécanismes relèvent de l'organisation de notre esprit, la *Grammaire Générale et Raisonnée*, pour sa part, noie cette distinction dans le chapitre qu'elle consacre aux pronoms relatifs, comme si, tout d'un coup, le résultat de l'analyse grammaticale était remis en cause en raison de son incapacité à satisfaire les exigences de l'analyse logique. Dès lors, si la proposition n'est que la formulation linguistique d'un jugement, il est possible de renverser l'ordre grammatical pour trouver une forme d'incidente dans la principale, ce qui est aujourd'hui déstabilisant pour nous, car le syllogisme qu'il faut entendre sous le masque de la grammaire est : *Dieu a créé le monde* [*Dieu est invisible*] + [*Le*

*monde est visible*], dans lequel se trouve « la forme démembrée d'une proposition complexe »<sup>3</sup>.

## 2. 4. De la fin du XVIII<sup>e</sup> siècle au XIX<sup>e</sup> siècle : de l'analyse des parties du discours à l'analyse de la phrase dans la première grammaire scolaire

À la fin du XVIII<sup>e</sup> siècle, sous l'influence exercée par le courant philosophique de l'Idéologie<sup>4</sup> qui supplante alors la tradition métaphysique, la grammaire voit se transformer son champ d'application et ses protocoles d'analyse. En voulant ramener chacun des éléments du discours à la classe qui lui est propre, en décomposant les formes de l'énoncé afin de parvenir à leur simplification, et en reprenant le principe cartésien de la réduction de la complexité par décomposition de ses parties, l'analyse grammaticale n'embrasse que les mots<sup>5</sup>.

Cette façon superficiellement modernisée de « faire les parties du discours » évacue totalement la question de la simplicité et de la complexité de la phrase.

En raison du mélange mal assimilé des principes d'une grammaire générale en désuétude et des objectifs orthographiques<sup>6</sup> de la première grammaire scolaire, les rapports de la simplicité et de la complexité de la phrase prennent encore le pas sur l'observation de l'existence de phrases simples et de phrases complexes, et finissent par se noyer dans l'opposition forcée que Landais introduit entre *construction analytique* (Sujet-Verbe-Complément) en référence aux principes de la grammaire générale :

La *construction analytique* est donc la seule nécessaire au langage. Les langues diffèrent dans les noms des différents objets ; dans les mots destinés à exprimer les qualités, les modifications, les attributs de ces objets, dans les accidents des mots ; dans la manière d'exprimer la corrélation ou la dépendance mutuelle des mots ; dans l'usage des métaphores et

---

3. Jean-Claude Pariente 1984 : 59

4. Antoine Destutt de Tracy 1796.

5. Jean-Edme Serreau & François-Narcisse Boussi 1824 : 370.

6. André Chervel 1977.

autres figures propres à chaque langue ; dans les tours de la *construction usuelle* ; dans les idiotismes. Mais, dans toutes, la pensée qu'on veut énoncer est représentée par des mots qui, dans la *construction analytique*, doivent avoir entre eux le même ordre, les mêmes rapports, qu'ont entre elles les idées partielles que chacun d'eux représente. (Landais 1834 : 397b)

et *construction figurée*, en référence au génie intrinsèque propre de chaque langue, en l'occurrence de la française, qui s'écarte de diverses manières de l'ordre analytique et simple :

La *construction figurée* est donc celle où l'on ne suit pas l'ordre et les procédés de la *construction analytique*, mais dans laquelle les procédés de cette dernière doivent pouvoir être toujours facilement aperçus, substitués ou suppléés. On l'appelle figurée, par ce qu'elle prend une forme, ou une *figure* qui n'est pas celle de la *construction analytique*. (*Ibid.* : 398b)

L'embaras évident de Landais met en lumière la contradiction interne du titre de son ouvrage : *Grammaire générale des grammaires françaises, présentant la solution analytique, raisonnée et logique de toutes les questions grammaticales anciennes et modernes*, et la confusion généralisée qui frappe la réflexion grammaticale et son enseignement par l'école au premier tiers du XIX<sup>e</sup> siècle. C'est donc bien autour de la notion de *figure* et la reconnaissance de la rhétoricité du langage en action, que se noue dans la première moitié du XIX<sup>e</sup> siècle la question de l'existence de phrases simples, syntaxiquement analysables [le mythe de la phrase S-V-C], et de phrases complexes, pour lesquelles il est nécessaire de recourir à des procédures d'analyse logiques.

Jusqu'à la fin du XVIII<sup>e</sup> siècle et les premières années du XIX<sup>e</sup>, la notion de *phrase* n'étant pas encore suffisamment stabilisée dans sa conception sémantico-syntactico-(typo)graphique, il était impossible que les adjectifs *incomplexe* /vs/ *complexe* puissent s'appliquer à elle. Ces termes ne pouvaient prédiquer, selon Noël et Chapsal, que la nature de compléments : « Il jouit du pouvoir » /vs/ « Il jouit du pouvoir du patron » (voir Bouard 2009).

## 2. 5. La 2<sup>e</sup> grammaire scolaire : la mise en place de l'analyse de la phrase complexe

Dans la deuxième grammaire scolaire, initiée à partir des années 1840-50, et qui se met définitivement en place dans les établissements scolaires après 1870, ultérieurement officialisée par la nomenclature de 1910, le point de gravité se déplace vers l'analyse logique de la phrase complexe. Mais ce déplacement n'est pas suffisant pour faire advenir une conception claire et stable de l'opposition *phrase simple* /vs/ *phrase complexe*, puisque le détour par la notion de *proposition*, en son sens logique, s'impose toujours comme en témoigne une grammaire comme celle d'Alexis Chassang en 1876 comme dans sa 8<sup>e</sup> édition de 1882 :

§ 157. Une *phrase* est une réunion de mots formant un sens complet : sa forme la plus simple est la *proposition*, qui est l'énonciation d'un fait ou d'un jugement.

§ 158. Une *phrase* peut se composer de plusieurs *propositions*.

Il y a, dans une *phrase*, autant de *propositions* qu'il s'y trouve de verbes à un mode personnel exprimés ou sous-entendus<sup>7</sup>.

Les *phrases* sont séparées entre elles par un point ou par deux points ; les *propositions*, par la virgule ou le point et virgule, à moins qu'elles ne soient liées par une conjonction ou un pronom conjonctif.

§ 159. Toute proposition se compose de trois termes : *sujet*, *verbe*, *attribut* [...].

§ 170. A la différence des langues grecque et latine, où l'ordre des mots, tout en étant soumis à de certaines règles, offre cependant une grande liberté et de nombreuses combinaisons, la langue française range en général les mots d'une manière uniforme : 1<sup>o</sup> en premier lieu vient le *nom sujet*, accompagné de l'adjectif qui le suit ou le précède ; 2<sup>o</sup> après le sujet vient le *verbe* ; 3<sup>o</sup> le verbe est suivi de ses *compléments*, d'abord (en général) du complément direct, puis des compléments indirects. Cet ordre est imposé à la langue française par l'absence des *flexions*, c'est-à-dire des *désinences casuelles* qui, dans les langues anciennes, distinguaient le sujet et les divers compléments. Mais, s'il a l'inconvénient d'être monotone et de ne pas se prêter aisément à tous les mouvements de la pensée, il a l'avantage d'être immédiatement saisissable pour l'es-

---

7. On retrouve évidemment ici la théorie de l'ellipse... contre laquelle Ferdinand Brunot, entre autres, s'est si fréquemment et violemment insurgé.

prité, et de faciliter le travail de l'analyse logique et de l'analyse grammaticale. [...].

§ 171. Les propositions sont *simples* ou *composées*.

On appelle *propositions simples* celles qui n'ont qu'un sujet, un verbe et un attribut, ces trois termes étant ou n'étant pas accompagnés de compléments. Ex. : « Ce temple est mon pays » (Racine). – « Les efforts de la violence ne peuvent affaiblir la vérité » (Pascal).

On appelle *propositions composées*, celles qui ont plus d'un sujet, d'un verbe et d'un attribut. Ex. : « La jalousie et l'émulation s'exercent sur le même objet » (La Bruyère) Deux sujets, deux propositions : 1° *La jalousie s'exerce...* ; 2° *L'émulation s'exerce*. « Charles I<sup>er</sup> était juste, modéré, magnanime, très instruit de ses affaires et des moyens de régner » (Bossuet). Quatre attributs, quatre propositions : 1° *Charles I<sup>er</sup> était juste* ; 2° *Charles I<sup>er</sup> était modéré* ; etc.

§172. On distingue encore *les propositions coordonnées* et *les propositions subordonnées* :

1° Ou bien les propositions sont mises l'une après l'autre, jointes ou non jointes entre elles par des conjonctions (*et, ou, ni, mais, or, donc, car, en effet, d'ailleurs*, etc.), de manière à rester indépendantes : ce sont des *propositions coordonnées* [...].

2° Ou bien une des propositions est dépendante de l'autre ; la proposition dépendante est dite *subordonnée*, et l'autre s'appelle *proposition principale*.

§ 173. Les *propositions subordonnées* les plus fréquentes sont celles qui commencent, soit par les conjonctions *que, si, lorsque*, soit par un pronom ou par un adjectif interrogatif ou conjonctif (*qui, quel, quelle*) [...].

REMARQUE 1. – Le vrai caractère de la *proposition subordonnée*, c'est de ne présenter un sens complet que par son union avec la *proposition principale*.

REMARQUE 2. – Une proposition subordonnée peut dépendre d'une autre proposition subordonnée.

REMARQUE 3. – On distingue trois espèces de *propositions subordonnées* :

1° On appelle *propositions complétives*, celles qui sont indispensables pour compléter le sens de la proposition principale, et qui lui sont rattachées par la conjonction *que*, dépendant d'un verbe précédent, ou par un mot interrogatif. Ex. : Je ne crois pas *qu'il soit de retour*. Je ne sais *comment répondre*...

2° On appelle *propositions circonstancielles*, celles qui ne font que modifier la proposition principale en y ajoutant des *circonstances* de cause, de

temps, de lieu, etc. Elles sont rattachées à la principale par les conjonctions *quand, lorsque, si*, etc. [...].

3° On appelle *propositions incidentes*, celles qui sont jointes à la principale par un pronom conjonctif et qui, par conséquent, lui sont moins étroitement unies que les *complétives*. Ex. : Celui 1° – *qui règne dans les cieux*, – 2° *et de qui relèvent tous les empires*, – 3° à qui seul appartient la gloire, la majesté et l'indépendance, est aussi le seul – 4° *qui se glorifie de faire la loi aux rois...* » (Bossuet).

REMARQUE –. On appelle *incise* une petite proposition, formant un sens complet, qui est intercalée dans une proposition plus étendue, et qui est mise soit entre parenthèses, soit entre deux virgules. [...] Il ne faut pas confondre une *incise* avec une proposition *incidente* [...]. (Chassang 1882 :181-182)

En dépit de la volonté didactique de l'auteur, il n'est nulle part ici fait allusion à une quelconque conception de la distinction entre *phrases simples* et *phrases complexes*.

À peine explorées par l'outillage rudimentaire du chapsalisme, les propositions à l'intérieur de la phrase complexe se voient dotées, dans la seconde moitié du XIX<sup>e</sup> siècle, d'une analyse calquée sur celle qu'on pratiquait dans le domaine de la proposition simple. Comme le remarque André Chervel :

Au terme de cette évolution [...] l'analyse logique du XX<sup>e</sup> siècle se préoccupe exclusivement du découpage de la phrase en propositions. (Chervel 1977 : 207)

Mais en propositions désormais soumises à des règles fonctionnelles d'organisation logique n'ayant plus qu'un lointain rapport avec les principes de la métaphysique dont leur dénomination était issue. Ce qui ouvre au XX<sup>e</sup> siècle sur la possibilité d'une autre conceptualisation.

### 3. L'invention d'une tradition

De l'émergence de l'analyse grammaticale des propositions au cours du XIX<sup>e</sup> siècle, jusqu'au milieu du XX<sup>e</sup> siècle, le couple conceptuel phrase simple/phrase complexe se met en place selon une terminologie qui varie d'une grammaire à l'autre.

### 3. 1. Le premier xx<sup>e</sup> siècle

La nomenclature officielle de 1910 ne mentionne pas les expressions. Le *Grand Larousse de la Langue Française* (1971-1977), au début de la partie lexicographique, mentionne Brunot (1922, *La Pensée et la langue*) pour l'opposition entre phrase simple et phrase complexe, sans que les syntagmes soient attestés dans cet ouvrage. Pour la première moitié du xx<sup>e</sup> siècle, nous disposons des résultats de l'enquête approfondie de Peter Lauwers (2004) dans un corpus de 27 grammaires du français, françaises et allemandes.

#### 3. 1. 1. Phrase simple

La terminologie « pour la proposition indépendante simple isolée témoigne d'un flottement plus ou moins généralisé » (266). Quatre paradigmes sont distingués :

- ◆ Phrase / proposition *simple* ou *isolée*
- ◆ Phrase / proposition *indépendante*

Ce terme est ambigu et cache en réalité trois découpages conceptuels différents :

- une proposition qui ne dépend pas d'une structure plus complexe et qui ne contient aucune proposition subordonnée
- une proposition qui ne dépend pas d'une structure plus complexe
- une proposition qui ne dépend pas d'une structure plus complexe, qui ne contient aucune proposition subordonnée et qui apparaît isolément (donc ni en coordination, ni en juxtaposition). (Lauwers 2004 : 267)

De ces trois découpages, c'est le premier qui domine<sup>8</sup>.

- ◆ Proposition / phrase *principale*  
« Dans un nombre considérable de grammaires, les propositions isolées sont rattachées à la catégorie des *principales* ». « On peut cependant supposer que ces auteurs savent faire la distinction entre une principale et une indépendante simple isolée ». (*Ibid.* : 267-268)
- ◆ Phrase / proposition

---

8. La deuxième conception est attestée chez Grevisse (1936) et Le Bidois (1935), la troisième chez Galichet (1947) et Cayrou (1948).

La terminologie est donc très variable, reflétant un flottement conceptuel qui affecte de même l'appréhension de la phrase complexe.

### 3. 1. 2. Phrase complexe

Elle reçoit un terme spécifique dans seulement 9 grammaires du corpus :

- *Phrase* (dans le sens de phrase complexe) : Académie (1932), Brunot
- *Phrase complexe* : Clédat, Galichet, Wartburg, Bruneau
- *Phrase composée* : de Boer (aussi *phrase entière*), Ulrix (1922) (aussi *groupe de propositions*), Grevisse.

Les grammaires qui sont dépourvues d'un terme propre rattachent le concept à la *phrase / Satz* en général (9 grammaires) ou s'en tiennent à des termes plus spécifiques, selon qu'il s'agit de cas de coordination ou de subordination (ou de simple juxtaposition). (Lauwers 2004 : 270)

Par exemple « *complex sentence* » correspond à la phrase à subordonnées, et « *multiple, double sentence* » à phrase à coordonnées / juxtaposées (Sonnenschein, cité p. 271)<sup>9</sup>.

La terminologie grammaticale devenue aujourd'hui traditionnelle apparaît donc encore en gestation dans la première moitié du siècle, comme en témoigne l'absence ou la concurrence des métatermes. La stabilisation terminologique date du second xx<sup>e</sup> siècle.

### 3. 2. Le second xx<sup>e</sup> siècle

Le couple terminologique n'est intégré à la nomenclature grammaticale officielle qu'en 1975. À partir des années 1960, la terminologie *simple / complexe* figure dans toutes les grammaires de référence<sup>10</sup>, à deux exceptions près :

---

9. Chez Damourette & Pichon (1911-1950 : 268-269)), « aucun terme spécifique n'est forgé pour la phrase complexe. Les auteurs réduisent tout à la notion de *phrase* à l'intérieur de laquelle les *sous-phrases* (= subordonnées) ne sont que des *compléments* [...]. On comprend que cette théorie se passe le plus souvent de proposition principale ; toutes les composantes sont ramenées à la phrase (intégrante) ».

10. Arrivé, Gadet & Galmiche (1986), Chevalier, Blanche-Benveniste, Arrivé & Peytard (1964), Denis & Sancier-Chateau (1994), Gary-Prieur (1985), Le Goffic (1993), Riegel, Pellat & Rioul (1994), Wagner & Pinchon (1969) et Wilmet (2010).

- ◆ Dans sa grammaire structurale, Dubois (1969) substitue au qualificatif *simple* l'adjectif *minimale* et ne mentionne pas la phrase complexe.
- ◆ Le Goffic (1993 : 11) y fait référence au détour d'un paragraphe, dans une parenthèse, sans l'intégrer au découpage de sa grammaire : « Quant à la catégorisation des subordonnées (qui était encore une question de Logique à l'époque classique, avec toute la problématique de la phrase complexe) ».

Si la phrase « complexe » ne recouvre encore que la phase à subordonnées dans la grammaire de Chevalier *et al.* (1964), l'indistinction avec la phrase à coordonnées et à juxtaposées se généralise par la suite.

Bien que cette tradition soit récente, le couple *phrase simple / complexe* est très tôt présenté comme traditionnel : ainsi en 1969 Wagner & Pinchon parlent déjà de « nomenclature traditionnelle » (1969 : 507).

### 3. 3. Le rendement conceptuel

Comment expliquer la généralisation des qualificatifs *simple / complexe* au détriment d'autres ?

Comme l'ensemble du vocabulaire de la grammaire traditionnelle, ces adjectifs ne sont pas spécifiques au métalangage, contrairement au terme *incomplexe* qui a pu émerger dans la grammaire générale. D'autre part, contrairement au couple *incomplexe/complexe*, qui pose comme norme de référence le pôle complexe, les termes *simple/complexe* se déterminent l'un par rapport à l'autre de manière réciproque, et à ce titre aussi apparaissent donc plus neutres.

La différence tient peut-être surtout au point de vue impliqué. Ainsi le terme *composé*, concurrent de *complexe*, qui présuppose l'opposé *élémentaire*, relève d'un point de vue plus logique que syntaxique (Arrivé 1986 : 531). Le couple *simple/complexe* fait quant à lui peut-être plus directement écho à l'opposition entre la phrase de l'enseignement primaire pour tous et la phrase de l'enseignement secondaire à destination des classes sociales favorisées, rapprochant la grammaire de la rhétorique – une partition qui s'est mise en place à partir de la fin du XIX<sup>e</sup> siècle (Balibar, Merlin & Tret 1974). Cette séparation est néanmoins schématique, car la forme sentencieuse et l'art de la pointe, qui ressortissent à la simplicité syntaxique, font

aussi partie de l'arsenal rhétorique. Rien n'est donc tranché tant du point de vue de l'analyse grammaticale que du point de vue de la valeur assignable aux différentes catégories.

## 4. Critique et solution

Le couple terminologique n'a guère été interrogé depuis son institutionnalisation, sans doute liée à un objectif pédagogique. Pourtant, si l'opposition relève bien de l'antonymie complémentaire (ou contradictoire), seul le second terme exprime une notion gradable : une phrase ne peut pas être plus ou moins simple, en revanche elle peut être plus ou moins complexe. Le couple présente donc une forme d'asymétrie problématique, en particulier la notion de phrase *complexe*, qui soulève au moins deux problèmes.

### 4. 1. Phrases complexes à subordonnées

Le couple terminologique *phrase simple/phrase complexe* est structurant dans la grammaire scientifique de Tesnière (1959), où le second terme fait l'objet d'un traitement original :

17. — Nous conviendrons d'appeler **phrase simple** toute phrase où l'agencement normal des nœuds n'est compliqué par aucun phénomène de jonction ni de translation.

18. — Inversement, nous conviendrons d'appeler **phrase complexe** toute phrase où interviennent à un titre quelconque les phénomènes de jonction ou de translation.

19. — Il est rare qu'une phrase ne contienne pas au moins un phénomène de jonction ou de translation. Aussi les **phrases complexes** sont-elles beaucoup **plus nombreuses** que les phrases simples.

20. — Néanmoins, comme il est de bonne pédagogie d'aller **du simple au complexe**, nous commencerons par l'étude de la phrase simple. (Tesnière 1959 : 101)

Le terme *complexe* recouvre donc les faits de *jonction* (juxtaposition et coordination) et les faits de *translation*, plus particulièrement la translation du second degré, qui équivaut à

la subordination<sup>11</sup>, notion dont l'originalité tient à une conception dynamique de changement de niveau syntaxique, sans trahir le souci didactique et pédagogique qui animait Tesnière.

Le point de vue intégrant de la grammaire générative, qui distingue phrase *matrices* vs *constituantes*, conduit à observer pareillement que les phrases complexes à subordonnées « n'ont pas une structure fondamentalement différente de celle des phrases simples » (Gary-Prieur 1985 : 40). Dans cette perspective, les propositions enchâssées (= subordonnées) assument la même fonction qu'un terme ou groupe non phrastique, ce qui constitue un paradoxe.

Mais la typologie *p. simple* /vs/ *p. complexe* pose un problème de nature historiographique dans la mesure où elle met sur le même plan, sous la catégorie « complexe », la parataxe et l'hypotaxe, alors que le degré d'implication n'est pas le même dans les deux cas. Or, au-delà d'une typologie problématique de la structure de la phrase, cette distinction hégémonique est partie prenante dans la difficulté à définir la notion de phrase qui caractérise la linguistique du xx<sup>e</sup> siècle. Elle constitue même l'un des points d'achoppement de sa définition. L'hétérogénéité de la catégorie « phrase complexe » s'ajoute ainsi à la difficulté d'une définition multicritères et à la notion de proposition héritée de la grammaire générale<sup>12</sup>.

## 4. 2. Les phrases complexes à propositions coordonnées et juxtaposées vs subordonnées : le traitement dans les grammaires

### 4. 2. 1. Riegel *et al.* (1994)

La question des types de phrases complexes (subordonnées, coordonnées, juxtaposées) n'est pas abordée par Le Goffic (1993), du fait que le grammairien raisonne à partir de l'oral, sans tenir compte de la segmentation typographique.

---

11. Il en existe trois types : la translation Verbe > Substantif (complétives), Verbe > Adjectif (relatives adjectives) et Verbe > Adverbe (circonstancielles).

12. Cf. Berrendonner 2017, article « Phrase », dans *L'EGF*.

La *GMF* est en revanche prolixe sur le sujet, mais reste doxique en dépit de l'amorce de critique (dans la première phrase) :

Rien de plus simple, du moins en apparence, qu'une **phrase complexe**. Syntaxiquement, une phrase est complexe si<sup>13</sup> [...] elle comprend un constituant qui, ayant lui-même la structure d'une phrase ( $P \rightarrow \mathbf{GN} + \mathbf{GV}$ ), se trouve ainsi être en relation de dépendance ou d'association avec une autre structure de phrase.<sup>14</sup>

[...] Les phrases complexes se distinguent traditionnellement selon leur mode de composition, c'est-à-dire selon la façon dont une ou plusieurs phrases constituantes [...], généralement appelées **propositions (2)**, s'insèrent dans la structure globale d'une phrase constituée dite *matrice*  $\mathbf{P}_0$  : (Riegel *et al.* 1994 : 469)

◆ **juxtaposition** [...]

◆ **coordination** [...]

◆ **subordination** [...]

◆ **insertion** (...) (Riegel *et al.* 1994 : 470)<sup>15</sup>

Cette grammaire consacre un assez long développement aux questions terminologiques :

Ce dédoublement terminologique [proposition = phrase constituante, proposition principale = phrase matrice] s'est étendu non seulement à la phrase simple appelée **proposition indépendante** (elle ne dépend pas d'une autre proposition et aucune autre proposition ne dépend d'elle), mais aussi aux phrases juxtaposées ou coordonnées également dites indépendantes parce que leur association ne porte pas atteinte à leur statut de phrases autonomes (elles n'entretiennent pas entre elles de rela-

---

13. Le premier critère est qu'« elle possède globalement les attributs définatoires de la phrase : elle a un type et l'unité mélodique correspondante, s'interprète comme une structure prédicative et peut constituer un énoncé complet » (Riegel, Pellat & Rioul 1994 : 469).

14. « Il s'agit là d'une manifestation du phénomène de la **récurtivité** à son plus haut niveau : à l'intérieur d'une phrase  $\mathbf{P}_0$  identifiée comme le niveau supérieur de construction syntaxique, on retrouve au moins une fois un constituant de la même catégorie ( $\mathbf{P}_1$ ,  $\mathbf{P}_2$ , etc.), selon le schéma simplifié de décomposition [...]. » (*Ibid.*)

15. « [...] les quatre types de constructions sont susceptibles d'exprimer les mêmes rapports entre propositions : il n'y a pas isomorphie entre les plans syntaxiques et sémantico-logique. » (*Ibid.* : 471)

tions de dépendance tout en formant ensemble une phrase complexe). (Riegel *et al.* 1994 : 471)

Dans l'analyse grammaticale de la phrase en propositions, on distinguera donc d'une part entre **phrases simples** (ou élémentaires) et **complexes** selon qu'elles comportent une ou plusieurs propositions ; d'autre part entre **propositions principales** et **propositions subordonnées** unies par un lien de dépendance orientée. Selon l'usage traditionnel, une proposition sera dite **indépendante**, si elle n'est pas subordonnée à une autre proposition et si elle n'inclut pas elle-même une subordonnée (ce qui permet de qualifier d'indépendantes les propositions coordonnées et juxtaposées en dépit des contraintes séquentielles qui régissent leurs cooccurrences). (Riegel *et al.* 1994 : 472)

La position de la *GMF* est conforme au paradigme dominant dans la première moitié du xx<sup>e</sup> siècle s'agissant de la phrase « indépendante » (cf. *supra* p. 8). Le qualificatif *simple* connaît pour concurrent l'adjectif *élémentaire*, tandis que le terme *complexe* ne connaît pas de rival, ce sur quoi l'on peut s'interroger.

Le critère syntaxique définissant la phrase comme un îlot de dépendances rectionnelles est mis à mal par les phrases à propositions coordonnées et juxtaposées. Cette disparité est pointée mais l'homogénéité de la catégorie « phrase » n'est pas questionnée :

Si l'on rapproche néanmoins à juste titre la juxtaposition et la coordination pour les opposer à la subordination, c'est qu'indépendamment de la présence ou de l'absence d'un terme de liaison, les deux premières opèrent sur le mode de l'**enchaînement parataxique** (qui joue également entre des mots et des syntagmes), alors que la troisième opère par **emboîtement hypotaxique** de propositions (à l'exclusion de tout autre constituant). Aussi les propositions juxtaposées et coordonnées gardent-elles leur autonomie syntaxique (et parfois énonciative), contrairement aux propositions subordonnées dont la forme est soumise à des contraintes syntaxiques, temporelles et modales imposées par l'élément dont elles dépendent. (Riegel *et al.* 1994 : 471)

#### 4. 2. 2. Wilmet (2010)

Dans sa grammaire « critique » du français, Wilmet pointe la variation terminologique :

La nomenclature varie d'un grammairien à l'autre. Flaux, p. ex., après avoir stipulé que « la phrase complexe (= à plus d'une proposition) s'oppose à la phrase simple (une seule proposition) » (1993 : 62), y range aussi bien les « propositions indépendantes » juxtaposées ou coordonnées (nos *phrases multiples*) que les « dépendantes » (nos *sous-phrases*). Goosse réserve *phrase complexe* « aux phrases contenant une ou plusieurs propositions » (1993, §212, b, 1°), mais il appelle *sous-phrases* les coordonnées « rattachées simplement l'une à l'autre » (ibid. 2°). Sa conception de la *phrase simple* (§212, a : « celle qui contient un seul prédicat ») assimilerait en réalité l'opposé *phrase complexe* à notre *phrase multiple*, vu qu'une *phrase matrice*, on le vérifiera, incorpore les *sous-phrases* à son sujet ou à son prédicat. (Wilmet 2010 : 636)

Il propose pour sa part un nouveau découpage conceptuel et terminologique. S'il ne conteste pas les subdivisions *phrase simple* et *phrase complexe*, il restreint néanmoins la signification de *complexe* à « comportant une *phrase matrice* et au moins une *sous-phrase* en hypotaxe », autrement dit phrase « à proposition(s) subordonnée(s) », comme le faisait Sonnenschein en 1912 :

Une phrase est *complexe* si elle se constitue de plusieurs mots réunissant une énonciation et un énoncé qui intègre au moins une autre phrase. La phrase intégrante peut être dite *phrase matrice* et la phrase intégrée *sous-phrase* (elle-même susceptible d'incorporer à la façon des poupées russes de nouvelles sous-phrases). (Wilmet 2010 : 636)

En complément il introduit la dichotomie entre « phrase *unique* (simple ou complexe) vs *phrase multiple* (additionnant des phrases simples ou complexes en parataxe) » (534), autonomisant de la sorte les phrases à propositions coordonnées et juxtaposées (ou insérées), celles-là mêmes qui posent problème pour la définition de la phrase. Il propose donc une étiquette synthétique et spécifique, « phrase multiple » :

Une phrase est *multiple* si elle se constitue de plusieurs phrases simples ou complexes associées par (1) emboîtement, (2) coordination, (3) juxtaposition. (Wilmet 2010 : 676)

Il semble résoudre ainsi le problème de la non-dépendance syntaxique des propositions de la phrase, sans renoncer à la notion de phrase.

Néanmoins l'expression « phrase multiple » existait déjà dans la grammaire de Sonnenschein (cf. *supra*). Au fond, s'agit-il d'autre

chose que d'un « relooking » terminologique, de l'introduction d'une spécification qui a le mérite d'attirer l'attention sur le problème, mais qui l'atténue plus qu'elle ne le règle ? La proposition de Wilmet consiste en effet en une typologie des phrases, et non en une unification de la catégorie. Témoignage de son intérêt pédagogique, elle a été reprise en 2018 par Cécile Narjoux, dans *Le Grevisse de l'étudiant. Grammaire graduelle du français*, qui vise le public des candidats aux concours de recrutement d'enseignants du secondaire, le CAPES et l'agrégation de lettres. Il semble donc que l'on s'achemine vers une vision tripartite de la phrase, simple, complexe ou multiple.

L'ouvrage que nous citons en introduction ajoutait :

Aborder la problématique *proposition-phrase-énoncé* [...] conduit toujours à une confrontation avec la difficulté terminologique. L'historiographie linguistique, ouverte à la diversité culturelle des traditions grammaticales, fait donc nécessairement porter sa réflexion sur la question terminologique et sur les problèmes traductionnels qu'elle est susceptible de poser. Les termes sont la vitrine lexicale d'une science, la surface très apparente de l'épistémologie qui la fonde. Parce qu'ils résultent de cette épistémologie, les termes portent avec eux tout le poids du discours de la science où ils sont engagés et requis. (Neveu 2019 : 6)

Dans le cas de la langue française, la distinction *phrase simple* /vs/ *phrase complexe*, qui recoupe globalement l'antique dichotomie entre *atticisme* et *asianisme* dans la rhétorique classique, est perceptible dès le XVII<sup>e</sup> siècle. Elle est apparue en liaison avec la naissance de la notion de subordination, dans le contexte d'un réseau notionnel autour de l'idée de *phrase*, laquelle commence à s'instituer tout au long de la seconde moitié du XVIII<sup>e</sup> siècle comme expression d'un énoncé au sens achevé.

Elle est ensuite présente dans la pensée de la grammaire scolaire au XIX<sup>e</sup> siècle, mais sans être précisément formalisée en tant que telle, les grammairiens employant toujours d'autres termes hérités de la tradition antérieure. Elle recoupe l'opposition entre analyse grammaticale et analyse logique à la faveur de l'opposition entre un enseignement primaire centré sur la phrase simple et un enseignement secondaire centré sur la phrase complexe, partition certes réductrice mais commode au plan pédagogique.

Très sporadique dans les grammaires de référence de la première moitié du XX<sup>e</sup> siècle, elle devient, dans la seconde moitié du siècle,

explicite dans les grammaires à orientation pratique et didactique, y compris dans les grammaires de référence, dans lesquelles elle joue un rôle structurant.

D'un point de vue scientifique, le couple *phrase simple/phrase complexe* apparaît néanmoins fragilisé par son orientation strictement pédagogique. Il n'est alors pas exclu qu'il remette la rhétoricité fondamentale du langage au premier plan sous couvert des grammaires du discours et des linguistiques de l'énonciation.

Ajoutons enfin qu'au-delà de ces flottements, nous n'aborderons pas ici la question du fondement épistémologique des différentes grammaires qui, au cours de l'histoire, se sont attachées à distinguer *proposition*, *phrase*, et *énoncé* avant de s'intéresser à ces satellites qu'en sont les notions de *phrase simple* et *phrase complexe*.

## Bibliographie

### Sources primaires

- Académie française. 1932. *Grammaire de l'Académie française*. Paris : Firmin-didot et Cie.
- Arnauld, Antoine & Claude Lancelot. 1803 [1660]. *Grammaire générale et raisonnée de Port-Royal*, précédée d'un *Essai sur l'origine et les progrès de la langue française*, par M. Claude Bernard Petitot, et suivie du commentaire de M. Charles Pinot Duclos, 2<sup>nd</sup>e édition. Paris : Bossange et Masson.
- Arrivé, Michel, Françoise Gadet & Michel Galmiche. 1986. *La Grammaire d'aujourd'hui. Guide alphabétique de linguistique française*. Paris : Flammarion.
- Beauzée, Nicolas. 1974 [1767]. *Grammaire générale ou exposition raisonnée des éléments nécessaires du langage, pour servir de fondement à l'étude de toutes les langues*. Paris : Barbou, Fac-simile reprint : Barrie E. Bartlett. Stuttgart-Bad Cannstatt: Friedrich Fromann Verlag (Günther Holzboog).
- Bruneau, Charles & Brunot, Ferdinand. 1933, *Précis de grammaire historique de la langue française*. Paris : Masson et C<sup>ie</sup>.
- Brunot, Ferdinand. 1922. *La Pensée et la Langue, Méthode, Principes et Plan d'une théorie nouvelle du langage appliquée au français*. Paris : Masson et C<sup>ie</sup>.
- Brunot, Ferdinand. 1932. *Observations sur la Grammaire de l'Académie française*. Paris : Librairie E. Droz.

- Cayrou, Gaston. 1948. *Grammaire française, Classes de grammaire*. Paris : Armand Colin.
- Chassang, Alexis. 1882 [1876]. *Nouvelle Grammaire française. Cours supérieur avec des notions sur l'histoire de la langue et en particulier les variations de la syntaxe du XVI<sup>e</sup> au XIX<sup>e</sup> siècle*, 8<sup>e</sup> éd. Paris : Garnier Frères.
- Chevalier, Jean-Claude, Claire Blanche-Benveniste, Michel Arrivé & Jean Peytard, avec la collaboration de Claudine Normand & Claude Régnier. 1964. *Grammaire Larousse du français contemporain*. Paris : Larousse.
- Clédat, Léon. 1894. *Grammaire raisonnée de la langue française*. Paris : Le Soudier.
- Damourette, Jacques & Édouard Pichon. 1911-1950. *Des mots à la pensée. Essai de grammaire de la langue française*, 7 vol. Paris : D'Artrey.
- De Boer, Cornelis. 1948. *Syntaxe du français moderne*. Leiden : Universitaire Pers.
- Denis, Delphine & Sancier-Chateau, Anne. 1994. *Grammaire du français*. Paris : Le Livre de Poche.
- Denys d'Halicarnasse. *Opuscules rhétoriques. Tome I. Les orateurs antiques*. Texte établi et traduit par Germaine Aujac. Paris : Les Belles Lettres (Collection Budé, 259). 1978.
- Destutt de Tracy, Antoine-Louis. 1796. *Mémoire sur la faculté de penser*. 2 floréal an IV. Paris : Mémoires de l'Institut national.
- Dubois, Jean. 1969. *Grammaire structurale du français*, t. 3 : *La Phrase et les transformations*. Paris : Larousse.
- Galichet, Georges. 1947. *Essai de grammaire psychologique*. Paris : PUF.
- Grevisse, Maurice. 1936 [12<sup>e</sup> éd. 1991], *Le Bon Usage*. Gembloux : Duculot.
- Gary-Prieur, Marie-Noëlle. 1985. *De la grammaire à la linguistique. L'étude de la phrase*. Paris : A. Colin.
- Jouvancy, Joseph de, Père. 1739. *Candidatus Rhetoricæ auctus, et meliori ordini digestus*. Paris : Barbou.
- Landais, Napoléon. 1835. *Grammaire de Napoléon Landais, Résumé général de toutes les grammaires françaises, présentant la solution analytique, raisonnée et logique de toutes les questions grammaticales anciennes et modernes*. Paris, Au Bureau Central, Rue du Faubourg Montmartre, n°15, A. Éverat, Imprimeur, Rue du Cadran, 16.
- Le Bidois, Georges & Robert Le Bidois. 1935. *Syntaxe du français moderne*. Paris : Picard.
- Le Goffic, Pierre. 1993. *Grammaire de la phrase française*. Paris : Hachette.
- Riegel, Martin, Jean-Christophe Pellat & René Rioul. 1994. *Grammaire méthodique du français*. Paris : PUF.

- Serreau, Jean-Edme & François-Narcisse Boussi. 1824 [1820]. *La Grammaire ramenée à ses principes naturels ou Traité de grammaire générale appliquée à la langue française*, 2<sup>nde</sup> éd. Paris : Dauthereau Libraire.
- Sonnenschein, Edward Adolf. 1912. *A New French Grammar. Based on the Recommendations of the Joint committee on Grammatical Terminology*. Oxford : Clarendon Press.
- Ulrix, Eugène. 1922 [1909]. *Grammaire classique de la langue française contemporaine*. Tongres : Vranken-Dommershausen.
- Wagner, Robert-Léon & Jacqueline Pinchon. 1969. *Grammaire du français classique et moderne*. Paris : Hachette.
- Wartburg, Walter von. 1934. *Évolution et structure de la langue française*. Berne : A. Francke
- Wilmet, Marc. 2010 [1997]. *Grammaire critique du français*. Louvain-la-Neuve : Duculot-Hachette Sup.

## Sources secondaires

- Andrieu, Wilfried. 2012. L'émergence de la syntaxe de la phrase complexe. *Vers une histoire générale de la grammaire française. Matériaux et perspectives*, dir par Jean-Marie Fournier & Valérie Raby. Paris : Honoré Champion. 781-796.
- Balibar, Renée, Geneviève Merlin & Gilles Tret. 1974. *Les Français fictifs : Le Rapport des styles littéraires au français national*. Paris : Hachette.
- Berrendonner, Alain. 2017. La notion de phrase. *Encyclopédie grammaticale du français* [[http://encyclogram.fr/notx/013/013\\_Notice.php](http://encyclogram.fr/notx/013/013_Notice.php), consulté le 29 janvier 2023].
- Blanche-Benveniste, Claire. 2013. Oralité et écriture : les unités de l'oral et de l'écrit. *Tranel (Travaux neuchâtelois de linguistique)* 58. 53-59.
- Bouard, Bérangère. 2009. La syntaxe dans la grammaire générale au milieu du XIX<sup>e</sup> siècle : structure de la proposition et transitivité. *Language and History* 52(1). 3-25.
- Chervel, André. 1977. *Et il fallut apprendre à écrire à tous les petits Français. Histoire de la grammaire scolaire*. Paris : Payot.
- Combettes, Bernard. 2011. Phrase et proposition. Histoire et évolution de deux notions grammaticales. *Le français aujourd'hui* 173(2). 11-20.
- Lauwers, Peter. 2004. *La Description du français entre la tradition grammaticale et la modernité linguistique. Étude historiographique et épistémologique de la grammaire française entre 1907 et 1948*. Louvain : Peeters.
- Miestamo, Matti. 2009. Implicational hierarchies and grammatical complexity, éd. par Geoffrey Sampson, David Gil & Peter Trudgill. *Language complexity as an evolving variable*. Oxford : Oxford University Press (*Studies in the Evolution of Language* 13). 80-97.

- Narjoux, Cécile. 2018. *Grammaire graduelle du français*. Louvain-la-Neuve : De Boeck Supérieur (Le Grevisse de l'étudiant. CAPES et agrégation lettres).
- Neveu, Franck, éd. 2019. *Proposition, phrase, énoncé : linguistique et philosophie*. Londres : Isteditions (Les concepts fondateurs de la philosophie du langage).
- Pariente, Jean-Claude. 1984. Grammaire et logique à Port-Royal. *Histoire Épistémologie Langage* 6(1) [Logique et grammaire]. 57-75. [[https://www.persee.fr/doc/hel\\_0750-8069\\_1984\\_num\\_6\\_1\\_1176](https://www.persee.fr/doc/hel_0750-8069_1984_num_6_1_1176), consulté le 29 janvier 2023].
- Raby, Valérie. 2002. L'analyse de la phrase complexe dans la grammaire générale. Construction des catégories « proposition principale » et « proposition subordonnée ». *Histoire Épistémologie Langage* 24(1). 93-105. [[https://www.persee.fr/doc/hel\\_0750-8069\\_2002\\_num\\_24\\_1\\_2852](https://www.persee.fr/doc/hel_0750-8069_2002_num_24_1_2852), consulté le 29 janvier 2023].
- Raby, Valérie. 2005. Proposition et jugement dans les *Annales de Grammaire* : le débat entre Michel de Neuville et Scott de Martinville. *Les Prolongements de la Grammaire Générale en France au XIX<sup>e</sup> siècle*, éd. par Jacques Bourquin. Besançon : Presses Universitaires de Franche-Comté. 135-150.
- Raby, Valérie. 2017. Points de vue sur l'énoncé et typologies propositionnelles dans la grammaire générale française (XVII<sup>e</sup>-XVIII<sup>e</sup> siècles). *Langages* 205. 103-116.
- Raby, Valérie. 2018. *Les Théories de l'énoncé dans la grammaire générale*. Lyon : ENS Éditions (Langages).
- Seguin, Jean-Pierre. 1993. *L'Invention de la phrase au XVIII<sup>e</sup> siècle : contribution à l'histoire du sentiment linguistique français*. Paris : Bibliothèque de l'Information grammaticale, Éditions Peeters.
- Seguin, Jean-Pierre. 2007. *La vraie révolution de la langue des Français*, Société Internationale de diachronie du Français [<https://diachronie.org/inédits/jean-pierre-seguin-la-vraie-revolution-de-la-langue-des-francais/>, consulté le 29 janvier 2023].
- Tesnière, Lucien. 1959. *Éléments de syntaxe structurale*. Paris : Klincksieck.



# Éloge de la naïveté dans les grammaires du français (xvi<sup>e</sup>-xvii<sup>e</sup> s.) Du naïf de la langue à la naïveté du style

Cendrine Pagani-Naudet<sup>1</sup>, Nathalie Fournier<sup>2</sup>

<sup>1</sup> Université Côte d'Azur, BCL UMR 7320 / HTL

<sup>2</sup> Université Lumière-Lyon 2, IRHIM UMR 5317 / HTL

---

## RÉSUMÉ

Dès les premières grammaires du français apparaissent des considérations relatives à la complexité et à la simplicité de la langue : tout en célébrant la richesse de sa langue, le grammairien doit démontrer sa capacité à être réduite en règles et dégager ce qui fait son ordre propre. L'enjeu est également d'ordre pédagogique : il s'agit de dédramatiser l'apprentissage d'une langue réputée difficile. Le travail du grammairien sera d'atteindre cet ordre qui fonde l'intercompréhension, et que les auteurs désignent aussi par le nom de *naïveté*.

La naïveté apparaît ainsi comme une notion mitoyenne, entre la simplicité et la complexité, qui définit l'objet et le but de l'activité grammaticale. Si la notion de naïveté endosse assez tôt une valeur esthétique, devenant une des perfections du style, elle désigne aussi le fonds commun de la langue, ce qui la caractérise et la distingue des autres idiomes. La recherche de la naïveté forme le socle de la norme d'usage.

Cet article étudie le parcours de cette notion au xvi<sup>e</sup> et au xvii<sup>e</sup> siècles, la manière dont elle se situe dans le discours sur la langue française, et participe à l'évaluation de la langue française au regard du latin, des autres vernaculaires et des variétés régionales.

## MOTS-CLÉS

naïveté, naturel, pureté, grammaire française, style

---

Pagani-Naudet, Cendrine & Nathalie Fournier. 2023. Éloge de la naïveté dans les grammaires du français (xvi<sup>e</sup>-xvii<sup>e</sup> s.). Du naïf de la langue à la naïveté du style. *Simplicité et complexité des langues dans l'histoire des théories linguistiques*, dir. par Chloé Laplantine, John E. Joseph & Émilie Aussant. Paris : SHESL (HELivres, 3).

**ABSTRACT**

As early as the beginning of French grammar in the 16<sup>th</sup> century, considerations can be found, concerning both the complexity and the simplicity of language. The grammarians have to celebrate the richness of their language and, at the same time, to demonstrate its capacity to be reduced into rules, so as to bring out its proper order. There is also a pedagogical matter at issue, to dedramatize French learning, which is commonly said to be an arduous matter. The grammarian's task will be to reach this order, which is the base for intercomprehension and is often called by the name *naïveté* (*naïveté*).

Naïveté appears as a middle notion, between simplicity and complexity, defining the aim of grammatical activity. Whereas the notion takes an early esthetic value, denoting one quality among style perfections, it still indicates the common core of language, its proper order, distinctive from other idioms. The search for naïveté constitutes the base for common use.

This paper deals with the evolution of the notion along 16<sup>th</sup> and 17<sup>th</sup> centuries, how it takes place in the discourse about French language, and takes part in the evaluation of French language, in comparison with Latin, other vernaculars and regional variations.

**KEYWORDS**

naïveté, natural, purity, French grammar, style

## 1. Introduction

Qu'est-ce qu'une langue *simple* ? Bien que le couple *simplicité/complexité* soit solidement établi dans le champ de la linguistique, les tentatives de conceptualisation se débarrassent difficilement des connotations que l'usage courant associe à ces termes. Aux critères formels qui permettent de caractériser et d'opposer langues simples et langues complexes, se superposent des enjeux idéologiques et culturels. Ces problématiques auxquelles entend répondre le présent volume ne sont pas étrangères aux grammairiens de la Renaissance et du début de l'Âge classique. Si les premiers grammairiens ne disposent pas des mots (*complexité* et *complexe* entrent tardivement dans le lexique), ils sont conscients de la complexité de leur langue, dont le « bâtiment », pour reprendre un terme de Louis Meigret<sup>1</sup>, est

1. Meigret ([1550] 1980 : 21).

un argument décisif en faveur de la « dignité » du français au regard du latin. Dans un contexte de concurrence et de hiérarchisation des langues, le grammairien doit affronter la variation régionale, le changement<sup>2</sup>, ou encore l'hétérogénéité introduite par les « escumeurs de latin », toutes choses qui donnent du français l'image d'une langue composite et instable. Il incombe donc au grammairien de prouver qu'un ordre existe. Il s'attache à rendre la langue *simple*, au sens où l'entendait Scaliger : mettre en évidence les règles permettant de ramener la diversité sous un seul chef<sup>3</sup>. Dans cette opération, le grammairien doit aussi donner de la langue qu'il décrit une représentation fidèle : il écrit pour permettre à un étranger de parler comme un authentique locuteur français. Il s'agit donc d'accéder au fonds de la langue, ce qui fait à la fois son ordre et son identité. Pour parler de cette double exigence, on pourrait presque dire de cette tension au cœur du travail grammatical, les premiers grammairiens parlaient du « naïf » de la langue.

Au XVI<sup>e</sup> siècle, le mot est plutôt pris en bonne part, dans ses emplois courants comme dans l'usage plus spécifique qu'en font les poètes et les rhétoriques. Son sens premier est celui de *nativus*, originel, et il constitue un doublet avec *natif*. Il renvoie aussi à ce qui est *naturel*, avec toutes les connotations positives qui se rattachent à l'idée de nature, d'authenticité des origines<sup>4</sup>. Ces valeurs sont à rapprocher de celles que véhiculent le mot *simple* si l'on en juge par des collocations qui associent les deux termes *simple* et *naïf* dont

---

2. Pour Charles Bovelles (1531), la variation régionale qu'il décrit rend impossible la mise en règle. Inversement, c'est le constat du changement rapide de la langue qui suscite l'appel de Geoffroy Tory à ce que « quelque Noble cuer semployast a mettre & ordonner par Reigle nostre Langage Francois » (Tory 1529, « Aux lecteurs »).

3. C'est un des mérites que Jean-Claude Chevalier signale au sujet de Scaliger : « un grammairien comme Scaliger définira donc très explicitement ce qui est, pour lui, la vertu d'une analyse grammaticale : c'est-à-dire la simplicité ; elle consiste à cerner sous un seul chef la diversité des caractéristiques à l'intérieur d'un genre qui sera le fondement de la construction grammaticale » (1967 : 25). Scaliger se dit en effet « guid[é] par la simplicité » (1540 : 353a) dans la description de la langue et veut imiter « la simplicité des anciens » (*ibid.* : 63b, 319b) ; nous remercions Bernard Colombat de nous avoir signalé ces passages.

4. Voir Le Cadet (2009 : 73).

la fréquence est moins significative que la célébrité de certaines d'entre elles (comme le « parler simple et naïf » de Montaigne<sup>5</sup>). Les deux termes ne sont pas synonymes mais se rapprochent du fait des connotations qu'ils mettent en commun, fondées sur un ensemble de valeurs morales, culturelles et esthétiques. *Simple* et *naïf* partagent ainsi l'idée de pureté<sup>6</sup>.

Les grammairiens vont s'emparer de ces différentes valeurs. Appliqué à la langue, *naïf* signifie ainsi « conforme à l'origine » dans un sens étymologique (cette conformité est rapportée au latin, comme chez Sylvius, où le *nativus* exhibe la parenté du français avec le latin) et historique (c'est le français des origines de la langue, « la langue de nos pères » d'Abel Mathieu, « le parler ancien » de Nicot). Seconde valeur de *naïf*, appliqué à la langue : le *naïf* signifie le *propre*, ce qui caractérise le français, ce qui lui est spécifique et le distingue des autres langues, anciennes et modernes (c'est la « naïve propriété de nostre parler » chez Maupas<sup>7</sup>). Cette propriété du français fait l'objet d'une double délimitation, externe au regard des autres langues, et interne, au regard des dialectes et aux variantes du français. Les grammairiens retrouvent ici un emploi que l'on rencontre aussi dans les rhétoriques lorsque les auteurs abordent la question de la traduction. Enfin, troisième ensemble de valeurs, et qui vaut autant pour le style que pour la langue, le terme se charge de connotations esthétiques, sociales et morales, le *naïf* est ce qui est conforme à la réalité, sincère, spontané, sans artifice. Au cours du xvii<sup>e</sup> siècle, le terme perd définitivement sa dimension linguistique : la naïveté devient une qualité du style et rejoint dans les grammaires

---

5. « Le parler que j'ayme, c'est un parler simple et naïf, tel sur le papier qu'à la bouche » (Montaigne 1580 : 231).

6. Voir ainsi les nombreuses collocations « simple et pur », « pur et simple » (« d'une espece tres especial aussi comme sont les parties de aer pur et simple et d'eau pure et simple », Nicole Oresme, 1377, *Le Livre du ciel et du monde*, cité d'après le *Dictionnaire du Moyen Français* [DMF]). Le lien entre *naïf* et *pur*, *naturel* est lui-même inscrit dans le lexique dès la période médiévale. La notice du *Dictionnaire du Moyen Français* [DMF] signale ainsi que le mot s'emploie pour un inanimé dans le sens de « naturel, brut ». On relèvera notamment son emploi pour qualifier un drap « dont la chaîne et la trame sont de qualité identique », ou encore la pierre qui n'a pas été taillée.

7. Maupas, 2021 [1618] : 182.

les remarques qui s’y rapportent. Comme le souligne Douglas Kibbee, c’est un mot à la mode, dont l’usage atteint un pic dans le dictionnaire de Nicot, mais qui ne dépassera pas le siècle, dans le domaine de la linguistique du moins (Kibbee 1993 : 175).

Nous nous proposons de revenir sur cette notion de « naïf François », concept qui permet aux grammairiens de construire et de représenter l’essence de la langue à partir des faits observables, marqués par la variation.

Nous suivrons un parcours chronologique, des années 1530 (la grammaire de Sylvius) jusqu’aux années 1660 (les grammaires de Chiflet et Irson, avant la *Grammaire générale*), articulé sur les trois grandes valeurs que nous venons de dégager : 1. Le naïf natif, originel et historique, 2. Le naïf propre à une langue, 3. Le naïf naturel, à valeur esthétique et morale. Nous suivrons également l’évolution du terme, qui quitte le champ strict de la grammaire pour fleurir comme terme stylistique chez les remarqueurs, à la suite de Vaugelas, qui affirme que la naïveté « est une des grandes perfections du style » (1647 : 487).

## 2. Le naïf de la langue au cœur de la « défense et illustration » de la langue française au XVI<sup>e</sup> siècle

Les grammaires françaises de la Renaissance comptent 34 occurrences de *naïf* et ses dérivés, avec de fortes disparités d’usage, signe peut-être de la fragilité du terme. Chez Palsgrave (3 occurrences), R. Estienne (1 occurrence) ou Cauchie (1 occurrence), *naïf* apparaît exclusivement dans les exemples. Ces derniers rendent bien compte du sens courant du terme : le sens de *naïf* recouvre à peu près celui de son doublet *natif*, tiré du latin *nativus*<sup>8</sup>, se rapprochant de l’idée de nature<sup>9</sup> et d’origine. La majorité des occurrences concerne toutefois

8. Voir par exemple les définitions données par Robert Estienne, dans son dictionnaire : « *Naïf* : Nativus, Genuinus, Germanus » (1549 : 399) et dans sa grammaire : « *Nativus* : *naïf* ou *natif* » (1557 : 96) ; ainsi que Cauchie : « *Naïfve, nativa* » (1586, f. 5r°).

9. « *Naturlyke* — m. *naïf* z, f. *naïfue* s; m. *genial* x, f. *genialle* s; m. *naturel* z, f. *naturelle* s (Palsgrave 1852 [1530] : 319), « These trees blome so kyndely that I

le métatexte. Meurier, Bosquet et Ramus emploient le terme de manière ponctuelle, et sans l'explicitier, au sujet de la prononciation : la « naïve prononciation » est l'objet de la description grammaticale et pour l'apprenant un modèle à atteindre<sup>10</sup>. Pour comprendre ce que recouvre la notion de *naïveté* à cette époque, c'est vers Meigret (14 occurrences) et Abel Mathieu (9 occurrences) qu'il faut se tourner. Eux seuls en effet en développent l'usage et l'inscrivent dans leur métalangage de manière variée pour qualifier la prononciation, une construction, une langue<sup>11</sup>, et notamment le français.

À ces disparités quantitatives vient s'ajouter une évolution sémantique : le sens de *naïf* se déplace « de l'idée d'origine, de qualité innée [...] vers des valeurs plus sociales – la simplicité, l'absence d'affectation, ou intellectuelles – la fidélité au réel » (Berlan 1997 : 49). Selon Françoise Berlan, ce glissement accompagne la succession de deux générations : l'une qui combat pour affirmer notre vulgaire face aux langues anciennes, l'autre qui affirme son caractère national face à l'italien envahissant.

## 2. 1. Retourner aux origines

Chez Sylvius, la grammaire étant rédigée en latin, ce sont les occurrences de *nativus* sur lesquelles ont porté les requêtes. La plupart des occurrences sont en rapport avec le son ou la prononciation, Colette Demaizière les traduit par *naturel*. Une seule, mais elle est significative, est traduite par *natif*. Elle se situe dans

---

truste we shall haue a good frute yere to yere : *Ces arbres germent or fleurissent si nayusement, que jespere que nous aurons des fruycts ceste année a planté* » (Palsgrave 1852 [1530] : 458).

10. Meurier intitule ainsi la section consacrée à la prononciation « Briefve instrvction, pour nayusement lire et prononcer le françois » (1557, f. 13v°). L'adverbe *naïvement* est inscrit dans le titre même de la grammaire de Bosquet, associé à l'idée de perfection : *Elemens, ov institvtions de la langve françoise, propres povr façonner la jeunesse, à parfaictement, et nayusement entendre, parler, et escrire icelle langue*, et dans la définition de la prosodie : « *Qu'est-ce de-Prosodie? Une partie de Grammaire, quy enseigne à nayusement prononcer, & distinctement lire* » (1586 : 135).

11. « Le pur et nayf Espagnol » (Matthieu 1559, f. 22r°), « il est aysé de distinguer le nayf françoys d'avecques le corrompu » (1559, f. 21v°).

la préface. Sylvius explique vouloir « décaprer la langue française et lui rendre son éclat *natif* », la « rétablir dans son antique pureté<sup>12</sup> » (trad. Demaizière : 200). La traduction par *natif* rend bien compte de ce désir de revenir à la naissance de la langue, à ses racines, et d'exhiber sa parenté avec les langues anciennes. Retrouver le *naïf* correspond à un mouvement de reconquête des origines<sup>13</sup>. Le travail du grammairien consiste alors à éliminer tout ce qui empêche d'y accéder. Au bout de cette quête, il y a bien un idéal de pureté et de simplicité. Ainsi au sujet des finales de la troisième personne, Sylvius observe la diversité des prononciations : les locuteurs ont multiplié les prononciations « corrompues », parce qu'ils « ne savent pas qu'il existe un système de leur langue ou s'ils savent qu'il existe, cependant ils ignorent totalement quel il est » (Sylvius 1531 : 119, trad. Demaizière : 334). Il s'engage donc à mettre à jour ce système et « à ce que la langue française soit le plus simple et pure possible, à l'imitation de la latine dont elle a découlé » (*ibid.*). Le retour au *naïf* (*nativum*) revient à rendre la langue *pure* et *simple* (*simplex et pura*<sup>14</sup>).

## 2. 2. Retrouver le principe de la langue

Ce retour à l'origine, pensé comme un retour aux langues grecque et latine, ne recouvre pas tout à fait la perspective de grammairiens comme Louis Meigret qui cherchent surtout à décrire le système en synchronie. Si les hommes se comprennent et échangent sur des sujets complexes, c'est que la langue « a en soi quelque ordre » et qu'on peut « la réduire à quelques règles » ([1550] 1980 : 1). De même qu'il induit de l'intercompréhension constatée l'existence d'un principe ordonnateur, Meigret va dans les différents domaines de la langue partir de ses observations, non pour les enregistrer et en subir l'extrême variété, mais pour atteindre les règles sous-jacentes.

---

12. « *Mihi vero ipse interim voti compos esse videbor, si nativum linguæ Gallicæ nitorem iamdiu propè exoletum, et situ obsitum, nom nihil detersero, ac velut postliminio in puritatis pristinæ partem restituero* » (1531 : 119).

13. C'est un mouvement que D. Kibbee a bien mis en évidence dans son étude sur Nicot ; voir *infra*, n. 27.

14. « [...] *vt lingua Gallica quàm maxime simplex sit & pura* » (Sylvius 1531 : 119).

Il remonte des effets aux causes. Cette manière de raisonner<sup>15</sup> est à l'œuvre lorsqu'il règle la morphologie verbale, gère et explique ses exceptions en sollicitant des formes profondes ou non attestées, de manière à rendre plus lisible le système<sup>16</sup>.

Mais ce travail s'effectue d'abord sur la prononciation et sa transcription, socle de toute la description grammaticale. La plupart des occurrences de *naïf* concernent ce domaine : dès le *Traité* de 1542, Meigret oppose la « commune écriture » à la « naïve prononciation française ». L'écriture « commune », surchargée de lettres inutiles, a corrompu la « naïve puissance des lettres ». Elle conduit à différentes lectures selon l'origine de celui qui oralise le texte, introduit le doute pour les étrangers, et entretient la variation régionale. Au grammairien donc de revenir à un système graphique simple par élimination des « lettres superflues », et de rendre aux lettres leur « naïve puissance » (Meigret [1550] 1980 : 15). Meigret jette bien un regard vers leur valeur d'origine : la puissance des lettres est « ancienne », héritée, il faut la garder et la restituer. Mais il atteint par là une sorte d'universel, ce fonds commun à toutes les langues, qui sont « forgé[e]s d'un même merrain de simples voix que nous appelons voyelles et consonantes » (*ibid.* : 6). Le système graphique qui en résulte est une reconstruction : la « naïve écriture » (*ibid.* : 11). Ramus dira « l'écriture grammairienne » (1572 : 55), ou « vraie écriture » (*ibid.* : 38). L'écriture est dite *naïve* parce qu'elle est univoque et reconduit sans ambiguïté vers la source orale. On lui attribue des valeurs de fidélité et de vérité.

Reste à saisir cette « naïve prononciation » (Meigret, [1550] 1980 : 41), et là les choses ne sont pas si claires. Meigret se fonde sur l'observation, sur « l'usage de la prononciation française » (*ibid.* : 18), mais de quel usage parle-t-il ? La « naïve prononciation » n'est pas disponible dans le réel. Meigret ne valide pas toutes les productions orales spontanées que peut saisir l'observateur. Il n'exalte pas les parlers régionaux, ne concède rien au Normand, et au « naïf Beauceron » (*ibid.* : 11). Le Parisien même n'est pas plus assuré d'avoir toujours son aval. Plus souvent, Meigret se réfère au

---

15. Sur cette démarche, voir Demonet 2021.

16. Voir Glatigny 1985.

courtisan (6 références), aux gens bien appris, qui réalisent un idéal de prononciation virile, qui fuit la mode et l'affectation<sup>17</sup>. Le *naïf* n'a donc plus rien du *natif*, du spontané, mais endosse déjà des valeurs morales et esthétiques.

En outre, quand l'usage branle, c'est au grammairien d'arbitrer, de sélectionner la meilleure variante, en se prononçant éventuellement contre l'usage observé. Meigret recommande ainsi de dire *liveau* et *livelet* et non *niveau* et *nivelet* comme disent les maçons. Il choisit les mots les plus proches de l'étymon source (« la naïve source »), et raisonne par analogie :

A bonne raison, donc, nous dirons *liueao*, et *liueler* : vu qu'ils sont descendus de *liure*, signifiant un certain poids : et que les dictions dont ils sont tirés sont *libella* et *librare*. [...] Il est donc manifeste que telles corruptèles de langage doivent être corrigées : et doit-on extirper ces façons de dictions qui entreprennent par erreur sur la naïve source que l'usage n'a pas du tout abandonnée. (Meigret [1550] 1980 : 102)

Le *naïf français* est ainsi une construction, résultat de l'activité rationnelle du grammairien qui ordonne pour mettre à jour un ordre fondamental. Cette activité est elle-même encadrée par l'oreille. Les règles formulées ne doivent pas être l'occasion de produire des formes indésirables. Intervient alors une oreille garante de la « naïve grâce française » (*ibid.* : 28) capable de contrôler la bonne formation des mots d'emprunt ou des énoncés. À cette oreille, Meigret prête un statut linguistique et collectif (Leclercq 2017). Toutefois, les connotations esthétiques ne sont pas absentes, de même que d'autres valeurs éthiques dont on investit la langue, porteuse d'un idéal social et culturel, qui s'exprime plus nettement dans la deuxième partie du siècle.

---

17. « Or, quant à moi, je ne poursuis pas ici cette douillette et efféminée façon de parler : car je la laisse aux amoureux, poursuivant tant seulement cette générale et commune façon qui sent son homme et qui est reçue entre les mieux appris. Or, que l'ε ouvert ne puisse être prononcé pour l'e clos, cette niaise prononciation que font aucuns des Parisiens [...] nous en donne une notable connaissance » (Meigret [1550] 1980 : 7).

## 2. 3. Observer ou reconstruire la naïveté du français

Dans les années 1560-1570, avec Ramus et Abel Mathieu, ou Henri Estienne, un autre mouvement se dessine. Il s'agit d'affirmer l'identité, la singularité du français. Ce mouvement se traduit par la glorification d'un passé strictement national et par la valorisation de traits distinctifs (francismes, régionalismes), qui ne peuvent être aisément réduits en règles.

La langue française a désormais un passé. Ramus affecte de parler de grammaire non pas « françoise » mais « gaulloise » (1572 : 1-2). Abel Matthieu<sup>18</sup>, son contemporain, se tourne vers « la langue de nos pères » (Matthieu 1560, f. 12v°). Il veut l'enrichir et l'« amplifier », en respectant sa « race première », « ce qu'elle a de propre en soi », sans la dénaturer. Cette identité de la langue s'affirme face à toutes les influences étrangères. Pour décrire « la langue patriote » (1559 : f. 5v°), il récuse l'art de la grammaire latine ou grecque (ce qui peut expliquer la singularité de son métalangage). La volonté de retrouver dans la langue un passé du terroir, indépendant des langues anciennes conduit à la célébration de la simplicité marotique<sup>19</sup>, et au culte des vieux auteurs français qui deviennent « un réservoir inépuisable d'expressions “propres” » (Lecointe 1993 : 532).

### 2. 3. 1. Où situer le naïf ?

Pour convaincus qu'ils soient de l'existence d'un français naïf, les auteurs de cette période ne s'accordent pas sur les lieux ou les milieux où il est possible de l'observer. Le modèle de Ramus semble être le peuple de Paris. Les lieux où il prétend trouver les manifestations

---

18. C'est, après Meigret, l'auteur qui présente les plus nombreuses occurrences de *nayf*, et c'est pour évoquer ces particularismes des vernaculaires : il évoque ainsi le « François nayf » (1559, f. 22r°) ou « nayf François » (*ibid.*, f. 21v°) mais aussi le « nayf Thoscan » (*ibid.*, f. 17r°), le « pur et nayf Espagnol » (*ibid.*, f. 22r°).

19. « s'accompent les Francoys qui composent ou devisent en la forme que jay dicté, et qui sont esloignez d'affectation et de motz à la moderne, entre lesquelz Clement Marot doit avoir place avecques les plus grandz du theatre, non pour son scavoir, ne pour son estude de grands authours, mais pour ce qu'il est vrayment amy de la multitude et qu'il ha la veine douce et coulante et l'esprit bon et facetieux » (Matthieu 1559, f. 21 r).

de cette langue gauloise sont bien connus, il ne s'agit pas des lieux hantés par les savants :

Lescolle de ceste doctrine nest point es auditoires des professeurs-Hebreus, Grecs, et Latins en luniuersite de Paris comme pensent ces beaux Etymologiseurs, elle est au Louure, au Palais, aux Halles, en Greue, a la place Maubert. (Ramus 1572 : 30)

Abel Matthieu pour sa part ne situe pas géographiquement la source du « François naïf » :

Quant au destroit de France à parler proprement, il est si court et si anguste qu'aujourd'huy on n'y sauroit asseoir le pied, ou le Francoys nayf y soit parlé et entendu du commun : mais il est repandu deca et dela ou sont les hommes bien appris, dont la plus part s'est retirée en la court du Roy, aux maisons des Princes et grandz seigneurs, ou es Justices souveraines et courtz de Parlement. (1559, f. 21v°)

Face à l'impossibilité de localiser l'excellence du parler français sur le territoire (contrairement à l'espagnol en Castille, ou à l'italien à Sienne<sup>20</sup>), Abel Matthieu s'oriente vers une référence sociale : le français *naïf* est observable chez les gens bien appris, autour du roi.

Ce que conteste Henri Estienne :

De quel François doncques enten-ie parler ? Du pur & simple n'ayant rien de ce fard, ni d'affectation : lequel monsieur le Courtisan n'a point encores changé à sa guise, & qui ne tient rien d'emprunt des langues modernes. (1569, préface)

Mais quant au lieu où l'on peut trouver ce français « bien entier et nayf » (1579 : 6), même s'il promet de le dire, en vérité Henri Estienne ne tranche pas. Il ne fait pas de Paris le lieu de référence absolu : « ainsi ne faut-il pas estimer que tout ce qui est du creu de Paris soit recevable parmi le pur et naïf langage François » (*ibid.* : 144). C'est que les manifestations de ce naïf sont diffuses par tout le territoire. Au grammairien de les saisir, y compris dans les dialectes.

---

20. « le pur et nayf Espagnol ha cours en Castille, le pur Italien quant au parler à Sienne, au proferer à Lucques, à l'escripre à Florence » (1559, f. 22v).

### 2. 3. 2. La richesse de la langue : idiotismes et régionalismes

H. Estienne affirme en effet la dignité des dialectes et montre la possibilité d'en enrichir la langue lorsque ces derniers possèdent un usage « plus beau ou plus significatif » (1579 : 187). Cela ne signifie pas qu'il accepte tout l'usage observé : qu'il s'agisse du langage parisien, ou des dialectes, il faut d'abord en « oster toutes les corruptions et depravations que luy fait le menu peuple », veiller à ajuster<sup>21</sup> le mot de manière à ce qu'il réponde à « la pureté françoise ». Etienne Pasquier, dans une lettre à M. de Querquifinen, seigneur d'Ardivilliers, aborde la question de « la vraie naïveté de notre langue et en quels lieux il la faut chercher » (1586, f. 51v<sup>o</sup>-54v<sup>o</sup>), et il adopte une position analogue à celle d'Henri Estienne. Il n'y a pas de lieu, pas de milieu où le français s'observe dans sa pureté. Il rejette la cour et le palais, mais invite à chercher le naïf partout où il se trouve, et à prendre aux dialectes tout ce qui permettra de trouver l'expression la plus énergique.

C'est le même désir de préserver la langue et ses richesses qui incite Ramus à revendiquer nos « francismes ». Entre 1562 et 1572, la revendication nationale aboutit à la reconnaissance d'idiotismes, anomalies qui ne rentrent pas tout à fait dans les cadres syntaxiques mais qui correspondent aux spécificités de la langue<sup>22</sup> et en sont les « ornements ».

Et si quelque Grammairien vouloit despouiller nostre langue de tels ornemens, Est ce moy? Est ce toy? Cest moy, Cest toy, ce seroit comme desgainer lespee luy tout seul a lencontre de toute la France. (Ramus 1572 : 168)

L'identité de la langue se lit dans un ensemble de traits irréductibles, notamment au moment de la traduction. Parmi les beautés de la langue, on relève *c'est moy* (qui avait d'abord été condamné par

---

21. Il y a en effet pour exprimer ce sentiment d'identité de la langue une image qui parcourt le siècle, celle du vêtement : il faut « accoutrer à la française » (Meigret [1550] 1980 : 27), « habill[er] » les mots « à la Françoise [...] pour les franchir et naturaliser » (Ramus 1572 : 66).

22. Le métatexte fait également l'objet de corrections, par exemple le passage de *est* à *c'est*. Voir Pagani-Naudet 2019.

Ramus en 1562). Autre trait que retient Ramus contre Meigret (et contre la raison) mais conformément à l'usage : l'accord du participe. L'évolution est nette entre une position plutôt critique en 1562<sup>23</sup>, et la reconnaissance en 1572 de l'usage, il donne « son suffrage à Marot » et se souvenant de la « souveraineté du peuple [...] reputé icy vng singulier Francisme » (1572 : 184).

Le français naïf n'existe nulle part, c'est une construction permanente, reposant non plus sur un principe de synthèse, de découverte d'un principe sous-jacent et ordonnateur (comme le pense Meigret) mais plutôt sur la sélection des meilleures variantes. Un idéal de pureté se joint à l'idée d'un enrichissement reposant sur l'accueil de tout ce qui, dans les parlers régionaux ou dans le passé de la langue, est compatible avec cet idéal. La naïveté est donc au xvi<sup>e</sup> siècle une notion complexe (et partagée entre grammaire et rhétorique), qui sert une vision militante, dynamique et esthétique de la langue. Le terme est polysémique : il conserve l'idée de pureté, de naturel, liée à son sens initial – naissance – mais se charge de connotations esthétiques (vérité, authenticité, simplicité).

### 3. Le xvii<sup>e</sup> siècle : de la naïveté de la langue à la naïveté du style

Au début du xvii<sup>e</sup> siècle, la notion de naïveté est encore vivante chez les grammairiens mais elle est en déclin, comme les termes qui y renvoient<sup>24</sup>.

Très éclairante pour l'évolution de la naïveté, est la comparaison entre Nicot, lexicographe et auteur du *Thésor de la langue françoise*,

---

23. « Partant, combien ce l'uzurpasion du pasif en tele transpozision soet alouer pour un bon plezir d'oreïes: neammoins m'amour vous e done, Dieu en se monde les a fet, e semblables syntaxes se pouroet débater, vu c'eles son' baties selon la vree rezon e regle de Gramere » (Ramus 1562 : 107).

24. On le constate par le nombre d'occurrences dans les grammaires : Masset 1606 : 1 ; Maupas 1618 : 7 ; Oudin 1640 : 0 ; Chiflet 1659 : 3 ; Irsou 1662 : 3 ; aucune occurrence par la suite dans le *Grand corpus des grammaires françaises*, Colombat *et al.* 2011).

*tant ancienne que moderne* (1606, David Douceur<sup>25</sup>) et Maupas, résolument grammairien et auteur de la *Grammaire & syntaxe françoise* (1607, 1618, 1625). Tous deux ont en effet une approche très différente du « naïf français », approche historique chez Nicot vs approche synchronique et intralinguistique chez Maupas. Tout aussi éclairante est la confrontation entre Maupas et Oudin, auteur d'une *Grammaire françoise rapportee à l'usage du temps* (1632, 1640), qui signe le tournant résolu vers l'examen des usages, tournant que prendront après lui Irsou (1656, 1662) et Chiflet (1659), sous l'influence décisive de la notion de « bon usage » et des *Remarques sur la langue française* de Vaugelas (1647).

On peut ainsi voir en Maupas le dernier grand témoin de la naïveté entendue comme notion opératoire appliquée à la description de la langue, et non du style<sup>26</sup>.

### 3. 1. Nicot et le « naïf François » comme français « natif »

D. Kibbee, dans un article très instructif de 1993, montre que Nicot inscrit son travail lexicographique dans le courant de ce qu'on peut appeler la « nouvelle historiographie », qui se développe en France dans la seconde moitié du xvi<sup>e</sup> siècle. Pour Nicot, écrit Kibbee, « l'histoire joue un rôle-clef dans la définition de ce qu'est le français »<sup>27</sup> » et le « naïf français » s'inscrit dans une filiation directe avec le français des origines, le français « natif ». L'entreprise lexicographique de Nicot est ainsi à la fois linguistique, morale et

---

25. Nicot commence par enrichir le *Dictionnaire François-latin* de Robert Estienne, puis donne son propre *Thresor*, qui paraît en 1606, deux ans après sa mort.

26. Chez Maupas (1618) : « la naïve propriété de nostre parler » ; « la naïve contexture & parole Françoise » (*Epistre*) ; « sans bonne et naïve prononciation » (f. 1v<sup>o</sup>) ; « la naïve & vraye prolotion de cette diphtongue [oi ou oy] » (f. 16r<sup>o</sup>), « l'antique & naïve [prononciation] » (f. 16v<sup>o</sup>) ; « la naïve difference d'entre Aux et és » (f. 35v<sup>o</sup>) ; « on dira plus naïvement un sçavant homme, un brave Soldat, [...] qu'au rebours » (f. 58v<sup>o</sup>).

27. « Nicot searches for the naïf français the direct line of language from the earliest French sources. [...] History plays a key role in the definition of Frenchness » (Kibbee 1993 : 168-169).

politique, selon un projet de restauration de la langue et des mœurs qu'annonce David Douceur dans son épître liminaire :

le restablisement de nostre parler ancien (plus ferme, plus court, et plus significatif, que celui qui a depuis esté receu) [est] le chemin de pouuoir reuenir à la generosité, constance, et magnanimité de nos peres. (in Nicot 1606)

Comme le montre Kibbee, que ce soit pour des questions de phonétique (*cap/chef*), de morphologie (*emperiere/imperatrice*), de concurrence entre un mot emprunté et un mot français (*accelerer/haster; embuscade/embusche*) ou de sémantique (*escumeur, rencontre*), Nicot défend toujours « le mot naïf François » contre les mots latins, dialectaux ou empruntés aux autres vulgaires, « la terminaison naïfve » contre « la terminaison estrangere », ou le sens naïf contre les « abusion[s]<sup>28</sup> ».

Ainsi, pour Nicot, le mot *cap* au sens de 'tête' « n'est pas naïf François », c'est un terme gascon ; en effet l'évolution phonétique du latin *caput* donne en français *chef* (« le François dit *chef*, ce que le Gascon dit *Cap* »). Le mot *emperiere* est le féminin français de *empereur*, qui se forme en ajoutant la « terminaison naïfue au masculin Empereur », alors que *imperatrice* est un emprunt au latin *imperatrix*, formé avec une « terminaison estrangere ». Dans le cas des doublets, Nicot défend systématiquement le mot français contre l'emprunt : *haster* « naïf François » et non *accelerer* « pur Latin » ; *embusche*, « naïf François » et non *embuscade*, « mot imité de la prononciation Espagnole *Embuscada* » ; *eschaugnette*, « mot naïf François » et non *sentinelle* « mot Italien naturalisé ». Quant au mot *rencontre*, il en maintient le sens originel d'événement fortuit (« ce que sans estre preueu, et inesperéement s'offre à nous ») et non pas délibérément produit (« ce qui s'offre avec pourchas »), ce qui est un sens « par abusion ». L'idée centrale de Nicot est, selon Kibbee, que « chaque forme de la langue (mot, morphème) a une valeur de base [*a core root value*], syntaxique, sémantique, morphologie, phonétique, que le temps a soit corrompue, soit laissée inchangée » (1993 : 169).

---

28. Voir les articles *Cap*, *Cheuvre*, *Empereur*, *Accelerer*, *Embuscade*, *Eschaugnette*, *Rencontre*.

Cette approche historique du naïf français chez Nicot a une conséquence décisive sur la définition du groupe de locuteurs de référence ; pour Nicot, ce ne peuvent être que les savants, les doctes, tels qu'on peut les trouver dans les cours de Parlement, mais ni la cour du Roi, ni Paris, et encore moins les provinces<sup>29</sup>.

Au XVII<sup>e</sup> siècle, les héritiers de cette approche historique sont Irson et son *Traité des Etymologies*, inclus dans sa *Nouvelle Méthode*<sup>30</sup>, et Ménage, qui, dans ses *Observations sur la langue françoise*, fort de son savoir étymologique, se met dans les pas de Ronsard<sup>31</sup> et défend « les antiques vocables [...] naïf[s] de la Langue Françoise », tels qu'ils sont préservés dans les Provinces :

Outre, je t'avertis de ne faire conscience de remettre en usage les antiques vocables : et principalement ceux du langage Vallon et Picard : lequel nous reste, par tant de siècles d'exemples, naïf de la Langue Françoise. L'entens de celle qui eut cours après que la Latine n'eut plus d'usage en nostre Gaule : et choisir les mots les plus preignans et significatifs, non seulement dudit langage, mais de toutes les Provinces de France, pour servir à la Poësie, lorsque tu en auras besoin. (Ménage 1676 : 162)

### 3. 2. Maupas et « la naïve propriété de nostre parler »

Lecteur assidu de Nicot, quoi qu'il en dise<sup>32</sup>, Maupas va adopter une position radicalement différente, synchronique et puissamment « theoricque » (1618, f. 136<sup>r</sup>°), ce qui le met du côté de l'entreprise rationnelle de Meigret, cherchant à rendre compte de l'« ordre » de la langue. Ce que Maupas entend montrer, sous la diversité des usages,

---

29. « *The choice of historical basis for 'true French' [...] leaves the determination of pure French to the scholars, rather than to the members of the court or the residents of Paris* » (Kibbee 1993 : 168-169) ; voir Trudeau 1992 : 85-86.

30. Ce traité est inclus dans l'ouvrage (1662 : 212-280), en dehors de la grammaire proprement dite, après une *Méthode pour composer des Lettres* (au sens d'épîtres) et un traité des *Regles generales de la Conversation*. Irson privilégie « l'origine prochaine des mots François » (213) dans une perspective de morphologie lexicale interne au français.

31. Ronsard : « Tu ne dedaigneras les vieux motz François, d'autant que je les estime tousjours en vigueur, quoy qu'on die » (*Art poétique François*, 1565 : 13).

32. Il se défend dans sa préface de toute influence métalinguistique, mais toutes ses listes lexicales témoignent bien de l'influence de Nicot.

c'est « la naïve propriété de nostre parler », c'est-à-dire de ce qui fait l'identité du français et le distingue des autres idiomes, latin et vernaculaires européens.

Cette perspective est explicitée dès l'*Epître* liminaire de l'édition de 1618. Maupas y explique comment il en est venu à écrire une grammaire pour lever les « difficultez » rencontrées par ses élèves étrangers (Maupas était maître de langues à Blois) :

Plusieurs aussi d'entr'eux [mes élèves] de bon esprit, me proposans les difficultez qui les travailloient, je les leur vuidois par une curieuse attention de la naïve propriété de nostre parler. (Maupas 2021 [1618] : 181)

On voit que la « naïve propriété » du français ne se confond pas avec la diversité des usages observables ; elle ne se dégage qu'au prix d'une « curieuse attention » de la part du grammairien, stimulé par les « difficultez » qui arrêtent les élèves étrangers. Maupas partage avec Meigret l'idée que décrire le « bâtiment de notre langage » (Meigret [1550] 1980 : 141), c'est rendre compte d'un système linguistique, dans lequel ce qui importe, ce n'est pas la diversité des faits d'usage, mais leur régularité (on peut les réduire en « reigles ») et leur propriété, c'est-à-dire leur spécificité par rapport à d'autres langues, le latin et les autres langues modernes. Maupas adopte ainsi ce que Jean-Marie Fournier et Valérie Raby appellent « une pratique différentielle des faits de langues vulgaires » (2014 : 339) qui est clairement explicitée au début du chapitre du Verbe<sup>33</sup> :

Traitant de l'usage des parties d'oraison, je me suis tous-jours proposé de toucher seulement ce qu'elles ont de particulière observation, autant qu'il me seroit possible, & que je m'en pourrois adviser, sans remplir le papier de reigles que nostre langue peut avoir communes avec la langue latine, ou autres vulgaires, ou qui autrement se peuvent suppleer par bon jugement. (f. 125r°)

### 3. 2. 1. La mise à l'écart du « naïf » historique

Maupas ne s'inscrit donc pas dans la perspective historique ouverte au XVI<sup>e</sup> siècle et reprise par Nicot. Sa grammaire est résolument

---

33. Ce principe sera rappelé à propos de l'adjectif (f. 57r°), et de la formation du passif (f. 123v°).

synchronique et il pourrait revendiquer, tout autant qu'Oudin, qu'elle soit « rapportée à l'usage du temps »<sup>34</sup>. Il n'envisage pas le français dans son évolution historique, ni à partir du latin<sup>35</sup>, ni à partir d'un français ancien, et qui serait valorisé en tant que tel. À aucun moment, sauf pour des questions de prononciation, la « naïveté » du français ne se définit comme l'héritage, préservé dans « nostre parler » actuel, d'un parler « antique », qui fonderait l'identité de la langue.

Maupas travaille ainsi dans une synchronie de principe, quasi-imperméable à l'archaïsme. Il ne relève que très sporadiquement des faits de langue qu'il juge « antique[s] » ou « anciens », et il n'en recommande pas l'usage. Ainsi de l'emploi de *Nulli* pour *personne*, dont il dit que « les anciens en usoient plus que ne faisons à cette heure » (f. 89r°) ou de « l'antique orthographe » de *cest*, pour *cet* (f. 69r°). La différence avec Nicot est évidente quand on regarde comment Maupas aborde la question féminin de *empereur* : il se borne à donner les trois formes « *Empereur, Imperatrice & Emperière* » (f. 41v°) et ne fait aucun commentaire sur la concurrence des deux féminins.

Le seul domaine, et encore dans quelques cas seulement, où le naïf puisse être assimilé à l'historique et recommandé en tant que tel, est celui de la prononciation. Ainsi, à propos de la diphtongue *oi* (dans les mots *foy, loy, Roy, voir, trois nois, &c.*), Maupas demande l'« antique & naïve » prononciation [wɛ], celle des « Doctes & bien-disans, és Cours de Parlement & ailleurs » et condamne la prononciation récente [ɛ] (comme dans les mots *mais, jamais, faire, plaisir*), en vogue « à la Cour du Roy » (f. 16 r°-v). C'est vraiment le seul cas où Maupas fonde explicitement la naïveté sur l'histoire et la charge d'une valeur positive et prescriptive. Les collocations sont d'ailleurs explicites : la prononciation [wɛ] est à la fois « antique & naïve », « naïve & vraye », et « vraye & ancienne » (1618, f. 16r°-v°) ; elle s'oppose à la « dépravation » et au « baragoin » (*ibid.*), et c'est, avertit-il, une condition de la « grace » : « la prononciation en toute

34. C'est le titre de la grammaire de Oudin (1632, 1640).

35. Le latin lui sert avant tout d'outil métalinguistique pour organiser les paradigmes et énoncer les règles morphologiques.

langue est dequoy l'on doit avoir plus de soin : car sans bonne et naïve prononciation, le langage perd toute sa grace » (f. 1v°).

### 3. 2. 2. La promotion du « naïf » comme principe interne et distinctif du français

Au naïf garanti par l'histoire, Maupas oppose le naïf interne au système de la langue, qu'il désigne dans l'*Épître* par « la naïve propriété de nostre parler ». Ce qui l'intéresse et qui fait pour lui le cœur de son activité de grammairien, c'est « ce que nostre langue a de propre » (f. 57v°). Cette naïveté de la langue va faire l'objet d'une double délimitation : d'une part, une délimitation qu'on peut dire externe, par rapport aux autres langues, latin et langues modernes, en écartant ce qui est « commun » pour mettre en évidence le « peculier » (f. 13r°), d'autre part une délimitation que l'on peut dire interne, par rapport à la diversité des variations dialectales ou sociolectales, qui de fait occupent peu de place dans l'ouvrage. La « naïve propriété de nostre parler » sert ainsi à construire et promouvoir un usage donné comme partagé, l'usage de « nostre langue », fondement de l'interlocution. On peut voir ainsi en Maupas un précurseur de la « langue commune » à laquelle l'Académie française dira « [s'être] retranchée » dans la préface de son *Dictionnaire* (1693).

Les questions phonétiques sont une bonne illustration de cette démarche. Maupas ne reprend pas la notion de « son naturel », c'est-à-dire le son des « lettres simples », voyelles et consonnes, qui constituent pour les grammairiens le fonds phonétique (sur le plan acoustique et articulatoire) des langues, « le merrain de simples voix », comme le disait Meigret ([1550] 1980 : 6). Certes, on peut comparer le son d'une lettre dans une langue à celui qu'elle a dans une autre langue ; ainsi le *A* français, « qui se prononce comme en Latin clairement à bouche ouverte », alors que les Anglais la « prononcent quasi comme nous la diphtongue *ai*, ce qui est de mauvaise grace en nostre langue » (f. 2v°). Mais ce qui retient avant tout Maupas, ce sont les sons qui sont propres au français, et à ce titre sont une difficulté pour les étrangers. Ainsi le *e* nasal, qui « devant *m*, ou *n*, se prononce vulgairement [i.e. ordinairement] quasi comme *a*. *Entendement, Emmener* » (f. 7r°) ; la prononciation

du digramme *GN* (la nasale palatovélaire [ɲ]), qui « donne un son qui ne peut estre bonnement déclaré par escrit », propre au français et dont les Allemands donnent une prononciation approximative : *onnon*, *minion* (f. 7v<sup>o</sup>-8r<sup>o</sup>) ; ou encore le *u* voyelle, dont le son [y] est « peculier à nostre langue » et que « les Allemans, Italiens & assez d'autres » prononcent [u] (f. 13r<sup>o</sup>).

Ces sons propres au français sont à extraire, le cas échéant, de la diversité des prononciations observables. Dans le domaine de la phonétique, Maupas ne tient pas registre des variations ; très peu sont relevées et, quand elles le sont, c'est pour être condamnées<sup>36</sup>. Dans le domaine de la phonétique, le souci de Maupas, qui hérite en cela de Meigret, est de réguler la prononciation. Ce qu'il qualifie de « bonne & naïve prononciation » est en fait une prononciation « modérée » (f. 11r<sup>o</sup>) et qui peut être commune à la diversité des locuteurs.

### 3. 2. 3. Une faible attention à la variation et à la diversité des usages

Maupas écrit sa grammaire pour un double public, pour les « estrangier[s] desirieux de nostre langue apprendre »<sup>37</sup>, et à ce titre c'est une grammaire de français langue étrangère, dans la filiation de Pillot, Cauchie et des méthodes d'apprentissage du français, mais aussi pour les « amateurs de la langue Françoise »<sup>38</sup>, ce qui l'autorise à proposer une « theoricque » poussée et des innovations, notamment pour l'article et le verbe. Dans cette double ambition, pratique et théorique, il n'y a pas de place pour la variation, pour la diversité des usages et la recherche du meilleur français dans des lieux ou des milieux particuliers, bref pour tout ce qui passionnait ses prédécesseurs de la fin du xvi<sup>e</sup> siècle. En tant que grammairien pour étrangers, Maupas privilégie des exemples courants, dans

---

36. Ainsi de la prononciation négligée du « bas populas », qui supprime le *ff* final dans *œuf*, *bœuf*, *soif* (f. 7v<sup>o</sup>), de « la niaise & vicieuse prononciation » de *Chouse pour Chose* (f. 9v<sup>o</sup>), et la condamnation sans appel de la prononciation en « *e* ouvert » de la diphtongue *oi* (f. 16r<sup>o</sup>-v<sup>o</sup>).

37. Sonnet liminaire.

38. Epistre dedicatoire.

des situations quotidiennes, tels qu'on peut les trouver dans les méthodes de langue ; et pour un théoricien, la variation fait plutôt écran lorsqu'on veut éclairer le système de la langue. Il est vrai que, et dans son ambition théorique, qui le rapproche de Meigret, et dans son peu d'intérêt pour le bon usage, Maupas peut apparaître dans les années 1620-1630, au moment où s'élabore la notion d'usage, comme un homme d'un autre âge. Mais n'oublions pas que ce « bloisien », ainsi qu'il se présente lui-même, est un provincial, qui est sans doute peu sorti de son périmètre natal et professionnel, que son contact avec l'aristocratie a été très limité<sup>39</sup> et que, à la différence de Oudin ou de Vaugelas<sup>40</sup>, il n'a pas fréquenté la cour et les milieux cultivés parisiens où l'on débattait de beau langage. Et puis, c'est un type d'esprit, un grammairien théorique, comme l'était Meigret et comme le seront Arnauld et Lancelot dans la *Grammaire générale et raisonnée* (1660).

### 3. 3. Oudin : de la « propriété » de la langue aux « impropriétés » des usages

Les choses changent avec Oudin, dont les préoccupations sont bien autres. Quand il présente en 1632 sa *Grammaire Française rapportée à l'usage du temps*, la naïveté n'est plus une notion opératoire pour saisir le fonds de la langue. Lui-même n'emploie d'ailleurs pas ce métalangage et ne s'interroge pas vraiment sur ce qu'est « la langue française » ni comment on pourrait la caractériser, par rapport au latin ou aux autres vernaculaires<sup>41</sup>. En 1632, il présente son ouvrage comme une grammaire « moderne » et une grammaire puriste,

---

39. Ce contact s'est limité aux grands seigneurs qui ont été ses élèves et peut-être même au seul grand seigneur qu'était le duc de Buckingham, auquel est dédié la 3e édition de 1625, parue après la mort de Maupas (voir N. Fournier 2021 : 10, 495).

40. Oudin est « Secretaire Interprete du Roy » (titre de 1632) et Vaugelas, savoyard mais issu d'un milieu cultivé, fréquente la cour et les salons dès son arrivée à Paris en 1601 (Ayres-Bennett 2018 : 7-14).

41. Il n'emploie l'expression qu'une seule fois dans un exemple *Je ne sçay pas la langue Francoise* (1640 : 238) ; cette attitude peut d'ailleurs surprendre de la part d'un italianisant et hispanisant, auteur de dictionnaires bilingues, français/italien et français/espagnol.

refondant celle de Maupas, qu'il juge à la fois archaïque, pédante et pleine d'erreurs :

Mon dessein n'estoit que d'augmenter la Grammaire du sieur Maupas : toute-fois y ayant recogneu force antiquailles à reformer, & beaucoup d'erreurs à reprendre, outre une confusion de discours repetez, obscurs, & pedantesques, je me suis resolu de vous en faire une moderne, afin de purger le monde en mesme temps des ordures que j'ay trouvées autre part & vous desabuser entierement. (*Épître Aux Curieux*, 1632<sup>42</sup>)

Cette épître au lecteur est significative à un double titre : Oudin s'y oppose en effet à la fois à Nicot et à Maupas.

Il s'oppose à Nicot, dans la mesure où sa condamnation répétée des « antiquailles » traduit son désir de s'émanciper du passé de la langue et de rompre avec la recherche de ses origines antiques et nationales. Ainsi pour lui, les « anciennes façons de parler »<sup>43</sup>, les « façons de parler un peu antiques »<sup>44</sup> comme les « vieux Auteurs, qui ont manqué à bien escrire »<sup>45</sup> sont suspects. Loin d'être valorisé comme authentique et propre au vrai français, ce qui est « ancien », « vieux », ou « antique », est stigmatisé comme « hors d'usage », « vitieux » ou « vulgaire »<sup>46</sup>. La condamnation de *à l'instar* ou de *moult* manifeste bien le lien entre condamnation (*point en usage*)

42. Cette épître n'est pas reprise dans l'édition de 1640.

43. « J'advertiray icy toutes sortes de personnes de ne se pas seruir des anciennes façons de parler ; comme celle-cy de Montagnes, *si i'auois des enfans ie leur desirasse*, au lieu de dire, *ie leur desirerois* » (1640 : 198).

44. « *sur ce il luy dit ; et pour ce, a cause de ce* ; où la particule *ce*, est mise pour cela : mais ie trouue ces façons de parler vn peu antiques » (1640 : 114).

45. « Et laissons les vieux Auteurs à part, qui ont manqué à bien escrire, faute de bien digerer, ou de rechercher la delicatesse, les phrases qui suivent le tesmoignent ; *quelque sien amy*, au lieu de, *quelqu'un de ses amis : vn certain leur parent*, pour vn certain de leurs parens : *quelque leur sujet*, au lieu de *quelqu'un de leurs sujets : aucun sien ennemy*, pour dire pas *vn de ses ennemis : vn autre mien valet*, pour *vn autre valet que i'auois : chaque sien effet*, pour *chacun de ses effets* : et si les possessifs en general se pouuoient construire en ce temps-là avec *quelque, aucun, certain, autre* et *chaque* : il n'est plus permis à cette heure [...] tout cela n'est plus receuable » (1640 : 121).

46. « *Finalemnt : finablement* antique et hors d'vsage » (1640 : 275) ; « Au demeurant, *doint, donray* et *donrois*, de *donner*, sont antiques et hors d'vsage, ou pour mieux dire vitieux » (1640 : 159) ; « *fors* et *fors* que sont antiques et

et référence historique, au passé antique (*tiré du latin*)<sup>47</sup> comme au passé national (*trop vieux*)<sup>48</sup>.

Mais c'est avant tout à Maupas que s'oppose Oudin, non seulement parce qu'il stigmatise ses « antiquailles » et son obscurité mais parce qu'il change radicalement de perspective. En effet alors que Maupas s'intéressait à la « propriété » de la langue, entendue comme un ensemble de traits distinctifs et systématiques, Oudin s'intéresse avant tout aux « improprietez » des usages, commises par différentes catégories de locuteurs fautifs – étrangers, provinciaux, commun peuple – qu'il oppose aux « vrais François »<sup>49</sup>, à ceux qui « parlent bien »<sup>50</sup>, aux « mieux disans »<sup>51</sup> :

Les estrangers mettent souuent *grand* pour *gros*, et c'est vne grande impropriété, parce que le mot de *grand* s'entend proprement des choses. (Oudin : 1640 : 87)

*Sortir* : *ie suis sorty*. Donnez-vous de garde d'vser de ce verbe en sens actif, pour *tirer dehors* ou *aeindre*, comme on fait en quelques Prouinces de France, où i'ay oüy dire, *sortez mon cheual de l'escurie*, *sortez cela de vostre coffre*, etc. Et tandis que je suis sur les improprietez, je vous en diray encore deux. [...] (1640 : 223-224)

Le commun peuple abusant de cette reigle, en discourant d'vne fille, dit, *c'est vne belle enfant*. I. vne belle fille ; *la pauure enfant*, etc. mais c'est vne grande impropriété. (1640 : 77)

*Desieuner*, nostre vulgaire fait improprement vn actif transitif de ce mot, lors qu'il dit, *il m'a desieuné de cét affaire-là*. I. il m'a entretenu ce matin de cela. (1640 : 211)

---

vulgaires » (1640 : 312) ; notons que le terme *vulgaire* est chez lui clairement péjoratif, il s'oppose aux habiles (1640 : 179).

47. « *A l'instar* est trop Latin, et n'est point en vsage parmy les bons François » (1640 : 293).

48. « *Moult* est trop vieux, et tiré du Latin » (1640 : 293).

49. « *dieutelet*, *arceau*, *bouueau*, *procuraceau*, *lionneau*, *hommeau*, *hommelet*, *cagnot*, *louueau*, qui sont tous diminutifs mal formez, et hors d'vsage parmy les vrais François » (1640 : 90).

50. « *Le vas*, et *vois*, pour la premiere personne, ne sont aucunement en vsage parmy ceux qui parlent bien » (1640, p. 160).

51. « Ces diphtongues *oi* et *oy* par vn vsage corrompu se prononcent diuersement : pour moy *ie me* rapporte aux mieux disans qu'elle doit estre leur meilleure prononciation » (1640 : 42).

La terminologie est significative. De la « propriété » de la langue aux « impropriétés » des usages, le passage métalinguistique se fait par un effet axiologique de péjoration. Au sens strict, une impropriété contrevient au caractère propre de la langue et des parties du discours<sup>52</sup> ; ainsi peut-on parler de l'emploi « impropre » d'un adverbe comme préposition, comme le fait Irson : « On dit proprement *il est hors de la chambre, ou dans la chambre*, et improprement, *il est dehors ou dedans la chambre* » (1662 : 97). De la disconvenance à la faute, le glissement est aisé et il est net chez Oudin : les impropriétés sont condamnées comme des fautes et le terme qui s'imposera est celui de barbarisme : « Faire vn *Barbarisme*, c'est vser d'vn mot qui n'est pas propre à la langue » ; c'est une faute contre la langue et contre « la Pureté du stile » (Chiflet 1659 : 147).

Chez Oudin donc, la conception de la langue change profondément : plus qu'à la langue, en tant qu'idiome, Oudin s'intéresse à la performance, aux usages « propres » ou « impropres ». Le tournant est pris : les grammairiens vont s'intéresser à la diversité des usages et se préoccuper de distinguer et recommander un bon usage. La naïveté n'est plus une notion que l'on peut appliquer à la langue, la langue elle-même se fragmentant en la diversité des usages ; elle est ainsi prête à glisser vers le style, comme marque de genre ou d'usage individuel.

### 3. 4. Grammairiens et remarqueurs, ou comment la naïveté passe de la langue au style

Une des dernières occurrences de la notion de naïveté appliquée à la langue se trouve chez Chiflet au début du traité de la prononciation, et *naïveté* y signifie bien le caractère propre, distinctif d'une langue :

Voicy vn Traité fort nécessaire, attendu que l'vne des plus grandes naïfuetez de chaque langue est en la prononciation. (Chiflet 1659 : 165)

---

52. Le terme *impropre* peut aussi marquer une dénomination métalinguistique inappropriée ; ainsi de l'appellation « *diphongue impropre* », récurrente dans les traités de prononciation et d'orthographe, qu'il faut entendre comme 'diphongue [i.e. digramme] improprement dite diphongue' ; on peut citer également l'article « impropre » *un* de Masset (1606 : 5), qui est un nom de nombre improprement dit 'article'.

C'est cependant avec le même Chiflet, que la notion de naïveté va passer dans le domaine du style, sous l'influence décisive de Vaugelas. Pour Chiflet, qui fait écho à Vaugelas (1647 : 487), la naïveté est une des « perfections du stile » (1659 : 147), et se caractérise par l'absence d'affectation et la fidélité au réel :

11. *La Naisveté*, [consiste] à refuir toute affectation, & à declarer sa pensée si naïvement, qu'il semble qu'on la voye à l'œil : deplus à l'artifice des descriptions, si naturelles, qu'elles soient comme vne peinture viuante de ce que l'on décrit. (1659 : 148)

Il en est de même pour Irson qui avertit d'user d'un « stile [...] tres simple et tres-naïf » dans les *Lettres familières* (1662 : 128) et recommande le stile « facile » de Voiture, qui tient à « une maniere de s'énoncer qui soit simple, naïve et familiere ; [...] tout y est naturel, et rien n'y est forcé » (1662 : 189).

C'est se faire l'écho explicite de Vaugelas, dont on sait l'influence décisive sur la réflexion métalinguistique du siècle classique. Vaugelas utilise souvent les termes *naïf*, *naïveté*, *naïvement*, mais toujours pour caractériser un fait de parole, et s'il lui arrive de parler de « la naïveté du langage » (1647 : 141), c'est pour caractériser le langage particulier d'un auteur (en l'occurrence Coeffeteau)<sup>53</sup>. Ce qu'il valorise c'est « la naïveté du style » (1647 : 487), la « naïveté de l'expression » ou de la « façon de parler » (1647 : 445). La naïveté prend ainsi une valeur à la fois sociale et esthétique et c'est une des manifestations de la civilité conversationnelle.

On peut se demander pourquoi les grammairiens comme Chiflet ou Irson font place dans leurs grammaires à l'entreprise descriptive et prescriptive de Vaugelas. La raison est que leur rapport à la langue a changé ; ce n'est plus du seul système de la langue dont il faut rendre compte mais de l'ensemble de ses usages. Chiflet s'en explique dans sa préface : ce que le grammairien a maintenant à tâche de décrire et d'enseigner, c'est « une Langue entierement formée, & qui est parvenuë au point de sa perfection : telle qu'est aujourd'hui la Langue Françoisse, fondée sur l'usage de la Cour, sur celuy des Maistres de la Langue, & sur celuy des bons Ecrivains » (Chiflet 1659, Préface). Son *Essay d'une parfaite Grammaire de la langue françoise* est un essai

53. Voir Ayres-Bennett 2018 : 403 (n. 551).

pour rendre compte d'une langue elle-même parfaite ; c'est donc une grammaire des usages et spécifiquement du bel usage, qu'il propose, et de ce bel usage, la référence incontournable est Vaugelas<sup>54</sup>.

La surprise vient de là où on ne l'attend pas, et c'est le père Bouhours qui fait briller l'ultime feu de « la naïveté de notre langue » dans le second des *Entretiens d'Ariste et d'Eugène* (1671), consacré à « la langue française ». L'éloge de la langue française, « langue polie et parfaite » (1671 : 39), passe par une comparaison à son entier avantage, avec la langue espagnole et la langue italienne. Entre « la pompe, & l'ostentation » de l'une et « la puérité & [l]es badineries » de l'autre, la langue française « garde un juste tempérament » et est « la plus simple & la plus naïve langue du monde » (1671 : 47). La naïveté de la langue, vue par Bouhours, par ailleurs auteur de *Remarques* et prescripteur de l'usage, est une notion militante, visant à assurer la prééminence du français sur ses concurrentes modernes, et dont le dessein politique est évident. La naïveté est une notion composite ; c'est d'une part la fidélité au réel et notamment à la pensée : la langue française « représent[e] naïvement tout ce qui se passe dans l'esprit »<sup>55</sup> (1671 : 50) et il n'y a qu'elle « qui suive la nature pas à pas » (1671 : 58) ; d'autre part, c'est le fondement de la civilité conversationnelle, dont les femmes « qui parlent naturellement et sans étude » (1671 : 59) sont le modèle, comme chez Vaugelas<sup>56</sup> ; enfin la naïveté est le propre du « beau langage » (1671 : 55), caractérisé en termes stylistiques :

Mais comme la langue française aime fort la naïveté [...] elle ne hait rien tant que l'affectation. Les termes trop recherchés, les phrases trop élégantes, les périodes même trop compassées lui sont insupportables. Tout ce qui sent l'étude ; tout ce qui a l'air de contrainte la choque. (1671 : 54)

---

54. On peut considérer que le programme de Chiflet correspond à ce que Deulofeu et Debaisieux (2012 : 27) appellent « une grammaire de référence », à la fois « somme organisée des connaissances » sur la langue et « recensement des usages effectifs [...] dans les divers registres », et il faudrait ajouter promotion d'un bon usage (cit. Pagani-Naudet 2021 : 86-88).

55. C'est notamment la célèbre question de « l'ordre naturel » (1671 : 57).

56. Vaugelas. *Préface des Remarques sur la langue française* (av<sup>o</sup>) ; voir Ayres-Bennett 2018 : 90-91.

## 4. Conclusion

Pour rendre compte de la langue française aux XVI<sup>e</sup> et XVII<sup>e</sup> siècles, de son unité comme langue (sa simplicité) et de la diversité, voire l'instabilité, de ses usages (sa complexité), les grammairiens ont répondu par la notion de *naïveté*. Empruntée à la rhétorique et notamment aux débats sur la traduction<sup>57</sup>, la notion peine à se fixer dans les grammaires, oscillant entre trois grandes valeurs : français « natif » (historique), français « propre » (systématique) et français « pur et simple » (variationnel). Entre 1530 et 1660, ces trois valeurs s'incarnent dans des filiations qui ne sont pas nécessairement contiguës chronologiquement. Nous avons ainsi tiré le fil historique du français naïf/natif de Sylvius à Nicot et Ménage, le fil systématique du français réglé par un « ordre », de Meigret à Maupas, et puis le fil variationnel du « pur & nayf François », qui mène de Ramus, Abel Mathieu, Henri Estienne jusqu'aux grammairiens comme Oudin, Irson et Chiflet, sous l'influence des remarqueurs, pour lesquels la naïveté est une des perfections du style.

Cette dernière acception l'emporte finalement, pour plusieurs raisons, à commencer par son ancienneté. La valorisation esthétique du naïf, présente dès le début, se maintient tout au long de la période sans solution de continuité, y compris chez les auteurs qui développaient une perspective linguistique de la naïveté : Sylvius l'associe à la pureté, Meigret et Maupas à la grâce. Deuxième élément de réponse constaté au fil de notre étude : la conception de la langue évolue, avec la perte de la dimension historique, amorcée avec Maupas et confirmée par Oudin<sup>58</sup>, ainsi que la prise en compte de la diversité des usages. À partir du moment où la langue n'est plus envisagée comme un système, ancré dans une synchronie de principe et transcendant la diversité des usages par le principe d'un parler commun – ce qui est la position de Meigret et Maupas – mais

---

57. Jacques Peletier recommande au « Translateur » de « garder la propriété et le naïf de la langue en laquelle il translate » (*Art poétique* [1555] 1990 : 245-246) ; voir Kibbee 2018 : 56-58.

58. Voir Brunot (1911 : 55-60). Il souligne l'ignorance de Vaugelas sur l'histoire de la langue, son incompréhension face aux mécanismes de changement qui sont à l'œuvre sous ses yeux sans qu'il en saisisse la nature.

doit se construire à partir des usages par une sélection des meilleures variantes et la détermination d'un lieu et d'un groupe de locuteurs de référence – position qui est celle de Ramus, Mathieu, Henri Estienne et de Vaugelas dans la notion de « bon usage » – la démarche va naturellement vers l'identification d'un groupe de référence, garant du *naïf* de la langue (les femmes, la Cour, les meilleurs écrivains). Enfin, au xvii<sup>e</sup> siècle, la question de l'identité de la langue ne se pose plus de manière aussi pressante qu'au xvi<sup>e</sup> siècle. Il s'agit désormais non plus de promouvoir la langue française mais de la perfectionner et de la fixer (hostilité à la bigarrure, à l'archaïsme, au néologisme). Le travail des grammairiens pour réduire la langue par règles et fonder son identité devient au xvii<sup>e</sup> siècle une ascèse stylistique (élimination de l'impur, de l'affectation). Mieux que le *naïf*, c'est le *simple* qui répond alors à cette approche, confondant sous un même vocable ce qui est désormais perçu comme une caractéristique du français<sup>59</sup> et une qualité du style<sup>60</sup>.

## Bibliographie

### Sources primaires

- Académie française. 1694. *Le dictionnaire de l'Académie françoise*/ Paris : Vve J. B. Coignard et J. B. Coignard.
- Arnauld, Antoine & Claude Lancelot. 1660. *Grammaire générale et raisonnée*. Paris : Pierre Le Petit.
- Arnauld, Antoine & Claude Lancelot. 2023. [1676]. *Grammaire générale et raisonnée*. Édition de Bernard Colombat et Jean-Marie Fournier. Paris : Classiques Garnier.
- Bosquet, Jean. 1586. *Elemens, ov institvtions de la langve françoise, propres povr façonner la jeunesse, à parfaitement, et nayement entendre, parler, et escrire icelle langue*. Mons : Ch. Michel.
- Bouhours, Dominique. 1671. *Les Entretiens d'Ariste et d'Eugène*. Paris : Sebastien Mabre-Cramoisy.

---

59. La Touche : « Comme la Langue Françoise est simple, et naturelle, il n'y a rien qui lui soit plus contraire qu'un style trop figuré. » (1730, t I : 341).

60. Irson : « Le [style] simple consiste seulement dans la façon de parler la plus naturelle et la plus ordinaire » (1662 : 128).

- Bouhours, Dominique. 1692 [1675]. *Remarques nouvelles sur la langue françoise*. Paris : George & Louïs Josse.
- Bovelles, Charles de. 1533. *Liber de differentia vulgarium linguarum et Gallici sermonis varietate*. Paris : Robert Estienne.
- Cauchie, Antoine. 1586. *Grammaticae gallicae libri tres*. Strasbourg : Bernard Jobin.
- Cauchie, Antoine. 2001 [1586]. *Grammaire française (1586)*. Traduction et notes de Colette Demaizière. Paris : Honoré Champion.
- Chiflet, Laurent. 1659. *Essay d'une parfaite grammaire de la langue françoise*. Anvers : Jacques Van Meurs.
- Chiflet, Laurent. 2021 [1659]. *Essay d'une parfaite grammaire de la langue françoise*. Édition de Cendrine Pagani-Naudet. Paris : Classiques Garnier.
- Dupleix, Scipion. 1651. *La Liberté de la langue Françoise dans sa pureté*. Paris : Denys Bechet.
- Dupleix, Scipion. 2018. *La Liberté de la langue Françoise dans sa pureté*. 1651. Edition de Douglas Kibbee et Marcus Keller. Paris : Classiques Garnier.
- Estienne, Henri. 1569 [1565]. *Traicté de la conformité du langage françoys*. Paris : J. Dupuis.
- Estienne, Henri. 1579. *De la précellence du langage françois*. Paris : Mamert Patisson.
- Estienne, Robert. 1549. *Dictionnaire françois latin*. Paris : Robert Estienne.
- Estienne, Robert. 1557. *Traicté de la grammaire françoise*. Genève : Robert Estienne.
- Irson, Claude. 1662 [1656]. *Nouvelle methode pour apprendre facilement les principes et la pureté de la langue françoise contenant plusieurs traitez*. Paris : Pierre Baudouin.
- La Touche, Pierre de. 1730 [1696]. *L'art de bien parler françois, qui comprend tout ce qui regarde la grammaire, & les façons de parler douteuses*, tome I. Amsterdam : Wetsteins & Smith.
- Masset, Jean. 1606. *Exact et tres-facile acheminement à la langue françoise, par Jean Masset, mis en latin par le mesme autheur pour le soulagement des étrangers*. Paris : David Douceur.
- Matthieu, Abel. 1559. *Devis de la langue françoise*. Paris : Richard Breton.
- Matthieu, Abel. 1560. *Second devis de la langue françoise*. Paris : Richard Breton.
- Maupas, Charles. 1618 [1607]. *Grammaire et syntaxe françoise*. Orléans : Boynard et Nyon.
- Maupas, Charles. 2021 [1618]. *Grammaire et syntaxe françoise*. Édition de Nathalie Fournier. Paris : Classiques Garnier.
- Meigret, Louis. 1542. *Traité touchant le commun usage de l'écriture françoise*, Paris, Denis Janot.

- Meigret, Louis. 1550. *Le Tretté de la grammere françoeeze*, Paris : Chrétien Wechel.
- Meigret, Louis. 1980 [1550]. *Le Traité de la Grammaire française*. Édition de Franz-Josef Hausmann. Tübingen : Gunter Narr Verlag.
- Ménage, Gilles. 1676. *Observations de Monsieur Ménage sur la langue française*. Paris : Claude Barbin.
- Meurier, Gabriel. 1557. *La Grammaire française*. Anvers : Plantin.
- Montaigne, Michel de. 1580. *Essais*. Bordeaux : Millanges.
- Nicot, Jean. 1606. *Thresor de la langue françoise*. Paris : David Douceur.
- Oudin, Antoine. 1640 [1632]. *Grammaire française rapportée au langage du temps*. Paris : Antoine de Sommaville.
- Palsgrave, John. 1852 [1530], *L'éclaircissement de la langue française*, Londres, Richard Pynson. Édition de Francis Génin. Paris : Imprimerie nationale.
- Pasquier, Etienne. 1586. *Les Lettres d'Estienne Pasquier*. Paris : L'Angelier.
- Peletier, Jacques. 1990 [1555]. *Art poétique. Traités de poétique et de rhétorique de la Renaissance*. Édition de Francis Goyet. Paris : Le Livre de Poche.
- Ramus, Petrus [Ramée, Pierre de la]. 1562. *Gramere*. Paris : André Wechel.
- Ramus, Petrus [Ramée, Pierre de la]. 1572. *Grammaire*. Paris : André Wechel.
- Ronsard, Pierre de. 1949 [1565]. *Abbrégé de l'Art poétique François*. Œuvres complètes, t. XIV. Édition de Paul Laumonier. Paris : Didier-Nizet.
- Scaliger, Jules César. 2018 [1540]. *De Causis linguae Latinae Libre trecedim*. Édition et traduction par Pierre Lardet, Geneviève Clérico et Bernard Colombat. Genève : Droz.
- Sylvius, Jacobus [Jacques Dubois]. 1531. *In linguam Gallicam isagoge et Grammatica Latino-Gallica*. Paris : Robert Estienne.
- Sylvius [Jacques Dubois]. 1998 [1531]. *In linguam Gallicam isagoge et Grammatica Latino-Gallica*. Édition et traduction de Colette Demaizière. Paris : Classiques Garnier.
- Tory, Geoffroy. 1529. *Le Champ Fleury*. Paris : Geoffroy Tory.
- Vaugelas, Claude Favre de. 1647. *Remarques sur la langue française utiles à ceux qui veulent bien parler et bien écrire*. Paris : Vve Jean Camusat.
- Vaugelas, Claude Favre de. 2018 [1647]. *Remarques sur la langue française utiles à ceux qui veulent bien parler et bien écrire*. Édition de Wendy Ayres-Bennett. Paris : Classiques Garnier.

## Sources secondaires

DMF : *Dictionnaire du Moyen Français*, version 2020 (DMF 2020). ATILF - CNRS & Université de Lorraine. [en ligne : <http://www.atilf.fr/dmf>].

- Ayres-Benett, Wendy. 2018. Claude Favre de Vaugelas. *Remarques sur la langue française de [1647]*. Édition critique de Ayres-Benett. Paris, Classiques Garnier.
- Berlan, Françoise. 1997. Langues naturelles et naturel des langues chez les théoriciens français d'Henri Estienne à Rivarol. *Nature et culture à l'âge classique (xvi<sup>e</sup>-xvii<sup>e</sup> siècles)*, dir. par Ch. Delmas et F. Gevrey. Toulouse : Presses universitaires du Mirail. 47-60.
- Brunot, Ferdinand. 1911. *Histoire de la langue française des origines à 1900. La formation de la langue classique*, t. III. Paris : Armand Colin.
- Chevalier, Jean-Claude. 1967. La Grammaire générale de Port-Royal et la critique moderne. *Langages* 7. 16-33.
- Colombat, Bernard, Jean-Marie Fournier & Wendy Ayres-Benett, dir. 2011. *Grand Corpus des grammaires françaises, des remarques et des traités sur la langue xiv<sup>e</sup>-xvii<sup>e</sup> siècles*. Paris : Classiques Garnier Numérique.
- Deulofeu Henri-José & Jeanne-Marie Debaisieux. 2012. Une tâche à accomplir pour la linguistique française du xxi<sup>e</sup> siècle : élaborer une grammaire des usages du français. *Langue française* 176. 27-46.
- Demonet, Marie-Luce. 2021. Le rythme de la prose chez Meigret. Des traductions au discours. *Actualités de Louis Meigret, humaniste et linguiste*, dir. par Véronique Montagne et Cendrine Pagani-Naudet. Paris : Classiques Garnier. 277-303.
- Fournier, Jean-Marie & Valérie Raby. 2014. Retour sur la grammatisation : l'extension de la grammaire latine et la description des langues vulgaires. *Penser l'histoire des savoirs linguistiques*, dir. par Sylvie Archaimbault, Jean-Marie Fournier & Valérie Raby. Lyon : ENS Éditions. 337-350
- Fournier, Nathalie. 2021. Introduction, in *Grammaire et syntaxe française de Charles Maupas (1618)*. Paris : Classiques Garnier.
- Glatigny, Michel. 1985. L'exception dans le système morphologique de L. Meigret. *Langue française* 66. 9-19.
- Kibbee, Douglas. 1993. Dictionaries and usage in 17<sup>th</sup>-century France. Le "naïf François". *History of Linguistics 1993*, dir. par Kurt R. Jankowsky. Amsterdam & Philadelphia : John Benjamins. 167-177.
- Kibbee, Douglas & Marcus Keller. 2018. Dupleix, *La Liberté de la langue Française dans sa pureté*. 1651. Édition de Douglas Kibbee et Marcus Keller. Paris : Classiques Garnier.
- Le Cadet, Nicolas. 2009. L'éloge de la naïveté dans les *Nouvelles récréations et joyeux devis*. Lire les « *Nouvelles récréations et joyeux devis* » de feu Bonaventure Des Périers, dir. par Dominique Bertrand. Clermont-Ferrand : Presses universitaires Blaise Pascal. 71-90.

- Leclercq, Odile. 2017. La «bonne oreille» dans le *Tretté de la grammere francoeze* de Louis Meigret. *Le jugement de l'oreille (xvi<sup>e</sup>-xviii<sup>e</sup> siècles)*, dir. par Agnès Steuckardt & Mathilde Thorel. Paris : Champion. 65-80.
- Lecointe, Jean. 1993. *L'Idéal et la différence. La perception de la personnalité littéraire à la Renaissance*. Genève : Droz.
- Pagani-Naudet, Cendrine. 2019. 1560-1572. Ramus et ses exemples. *Grammaticalia. Hommage à Bernard Colombat*, dir. par Jean-Marie Fournier, Valérie Raby & Aimée Lahaussais. Lyon : ENS Éditions.
- Trudeau, Danielle. 1992. *Les Inventeurs du bon usage (1529-1647)*. Paris : Éditions de Minuit.

# Hebrew as “lingua simplicissima”: complexity in orations on the Hebrew language at the sixteenth- century Louvain Collegium Trilingue<sup>1</sup>

**Maxime Maleux**

Research Foundation–Flanders (FWO) & KU Leuven

---

## ABSTRACT

In the sixteenth century, Hebrew became part of the linguistic horizon of humanists, being taught in university curricula. Yet, professors of Hebrew felt the need to defend the study of the language from critics, using a varied array of arguments. One of these arguments was the language’s simplicity, contrary to its reputation as a difficult language. This article investigates how these orations dealt with the language’s simplicity. An interesting case is an anonymous speech delivered in the context of the Hebrew lessons of the Louvain Collegium Trilingue. Its author describes the character of Hebrew as “simple yet perfect”. Here, we see that the complexity of the language is strictly connected to the theological aspect.

## KEYWORDS

Hebrew, humanism, orations, Trilingual College, Louvain

---

1. I would like to thank Raf Van Rooy and the two anonymous reviewers for their valuable criticisms and suggestions. The research for this paper was funded by the Research Foundation – Flanders (FWO), grant number 1145321N.

Maleux, Maxime. 2023. Hebrew as “lingua simplicissima”: complexity in orations on the Hebrew language at the sixteenth-century Louvain Collegium Trilingue. *Simplicité et complexité des langues dans l’histoire des théories linguistiques*, dir. par Chloé Laplantine, John E. Joseph & Émilie Aussant. Paris: SHESL (HEL Livres, 3). 137-152. 

**RÉSUMÉ**

Au seizième siècle, l'hébreu entra dans l'horizon linguistique des humanistes et fut enseigné dans les universités. Cependant, les professeurs d'hébreu se sentirent obligés de défendre l'étude de cette langue contre les critiques, en utilisant toute une série d'arguments variés. Un de ces arguments était la simplicité de l'hébreu, ce qui ne cadre pas avec l'image que l'on avait généralement de cette langue. Cet article vise à explorer comment ces discours traitent la simplicité de la langue. Un cas intéressant à ce propos est un discours anonyme délivré à l'université de Louvain, au Collège des Trois Langues. L'auteur y présente l'hébreu comme langue « simple mais parfaite ». Dans ce cas précis, la complexité de la langue est étroitement liée à son aspect théologique.

**MOTS-CLÉS**

hébreu, humanisme, discours, Collège des Trois Langues, Louvain

## 1. Introduction: a new language on the linguistic horizon of the humanists

In his 125<sup>th</sup> letter, addressed to Rusticus, St. Jerome (c.345-420) confessed that learning Hebrew had been a particularly difficult task for him. In addition, he characterized the language's aesthetic qualities in utterly negative terms<sup>2</sup>:

*12. Dum essem iuvenis et solitudinis me deserta vallarent, incentiva vitiorum ardoremque naturae ferre non poteram; quae cum crebris ieiuniis frangerem, mens tamen cogitationibus aestuabat. Ad quam edomandam cuidam fratri, qui ex Hebraeis crediderat, me in disciplinam dedi, ut post Quintiliani acumina Ciceronisque fluvios gravitatemque Frontonis et lenitatem Plinii alphabetum discerem, stridentia anhelantiaque verba meditarer. Quid ibi laboris insumpserim, quid sustinuerim difficultatis, quotiens desperaverim quotiensque cessaverim et contentione discendi rursus inceperim, testis est conscientia tam mea, qui passus sum, quam eorum, qui mecum duxere vitam. Et gratias ago domino, quod de amaro semine litterarum dulces fructus capio.*

'12. When I was a young man, though I was protected by the rampart of the lonely desert, I could not endure against the promptings of sin

2. On Jerome's views on and appreciation of Hebrew, see Denecker (2015) and Barr (1967).

and the ardent heat of my nature. I tried to crush them by frequent fasting, but my mind was always in a turmoil of imagination. To subdue it I put myself in the hands of one of the brethren who had been a Hebrew before his conversion, and asked him to teach me his language. Thus, after having studied the pointed style of Quintilian, the fluency of Cicero, the weightiness of Fronto, and the gentleness of Pliny, I now began to learn the alphabet again and practice harsh and guttural words. What efforts I spent on that task, what difficulties I had to face, how often I despaired, how often I gave up and then in my eagerness to learn began again, my own knowledge can witness from personal experience and those can testify who were then living with me. I thank the Lord that from a bitter seed of learning I am now plucking sweet fruits<sup>3</sup>.

Nevertheless, he expressed his pleasure at having persevered, because his knowledge of the holy language enabled him to accomplish one of the most important translations in history: the Latin *Vulgata*. Ironically, his translation caused most Western scholars to feel no need to learn Hebrew in order to read the original text of the Old Testament, and thus, for a long time, scholars maintained the same negative views of Hebrew as a language (in addition to its negative association with the despised Jews).

Before the advent of humanism in western Europe, linguistic horizons were rather limited; a medieval scholar needed to know only scholastic Latin to have access to scholarly literature. A few learned men, like William of Moerbeke (c.1215-1286), translated some Greek works, but they were the exception. The Holy Writ was only read in Jerome's Latin translation, which was unquestioned<sup>4</sup>. However, with the rise of the *studia humanitatis*, a desire to know the languages of the classical sources (encapsulated in the Latin saying *ad fontes*) led to the appearance of Greek (the language of the ancient pagan Greek literature and of the New Testament) and Hebrew (first and foremost the language of the Old Testament) in the universities' curricula. This expansion, however, was not easily achieved; when the first university-linked institution to learn Hebrew, the *Collegium Trilingue*

---

3. Translated by Frederic A. Wright (1933). All other translations in this article are my own.

4. For the interaction of medieval scholars with Jewish culture and scholarship, see Dahan (1990).

*Lovaniense*<sup>5</sup>, was established in 1517 with the help of Desiderius Erasmus (1466/67/69-1536), the study of Hebrew and Greek was immediately attacked by theologians who regarded the study of the biblical languages as a threat to their authority in Biblical exegesis. One of the most fervent opponents of the new institution was the theologian Jacobus Latomus (c.1475-1544), who wrote a pamphlet on the uselessness of learning languages other than Latin for serious theological study.

Since its beginnings, professors of the Collegium Trilingue (or college of the three languages) wrote Hebrew grammars for their students' use. Some of these became quite successful and popular, like Clenardus' *Tabulae in grammaticen Hebraeam* (Louvain, 1529)<sup>6</sup>. Since the Hebrew language is not related to Latin or Greek and has its own grammatical tradition (modeled on Arabic grammar), it was a major challenge for these grammarians to describe the language using a Latin(ate) terminology and grammatical framework.

The study of Hebrew also faced the anti-Judaism of most Western scholars. As the language of the Jews, it was often disparaged. Most apologists for Hebrew explicitly note that they praised Hebrew "without praising the Jews". For the teaching of Hebrew, recourse was usually made to Jews who had converted to Christianity, yet these new Christians were not free from criticism<sup>7</sup>. The strangeness

---

5. On the history of the Collegium Trilingue, see De Vocht (1951-55). Papy (2017) and Van Hecke (2018) constitute the first serious endeavors into the Hebrew teaching methods at the Trilingual College.

6. From the early days of the Collegium Trilingue, Hebrew books were printed in Louvain. The Hebrew script was fairly challenging for the recently developed technique of printing, but as the need for printed Hebrew grammars and editions rose, the local printing press of Thierry Martens (1446/7-1534) invested in Hebrew type. As such, he was able to print the necessary textbooks for the Trilingue's Hebrew classes, like, for instance, Clenardus' *Tabulae*. These *Tabulae* were reprinted thirteen times in France by Christian Wechel and annotated by illustrious Hebraists like Johannes Quinquarboreus (Jean Cinquarbres, c.1520-1565) and Johannes Isaac Levita (1515-1577).

7. One example from the Louvain Collegium Trilingue is Johannes Isaac Levita, who had been a rabbi before his conversion. He taught Hebrew together with Andreas Balenus and used Jewish commentaries in his exegetical treatises. This use of Jewish sources was criticized by his former student, the Bishop of Roermond, Wilhelmus Lindanus (1525-1588). Levita reacted by writing his

and difficulty of the language, with its guttural sounds, and its writing system (a consonantal alphabet written from right to left), make it relatively hard for beginners to learn; these features were frequently used as fodder for attacks on the language<sup>8</sup>. These are the points that apologists for the language refute in their orations.

In this article, I will concentrate on one such oration written in the context of the Louvain Collegium Trilingue. Among many other arguments, the oration explicitly mentions the language’s simplicity as an important feature. Was this argument made simply to attract new students? Or does this quality fit the humanists’ understanding of Hebrew grammar? Or is it perhaps a mere theological argument? In this case study, I will look at how these texts written by professors of the Louvain Collegium Trilingue relate to other documents within the same genre as regards the argument of simplicity.

## 2. Linguistic description of Hebrew in orations: the genre and practices in Louvain

Numerous orations promoting the study of languages like Greek and Hebrew appear in the sixteenth century and beyond. Especially at the dawn of this period, Hebraists felt the need to defend their object of study. A first notable example is the oration by Petrus Mosellanus (1493-1524), *De variarum linguarum cognitione paranda oratio* (Leipzig, 1518), in which he praised Erasmus for having established the Collegium Trilingue and advocated the study of languages as an indispensable skill for theologians<sup>9</sup>. In the seventeenth century, such texts continued to appear, but these later orations tended to focus more on academic questions regarding the provenance of Hebrew and

---

*Defensio veritatis Hebraicae sacrarum literarum* (1559), in which he defended the traditional Biblical scholarship by the Jews. See Dunkelgrün (2017: 418).

8. Vitus Winshemius in 1549 (fol. Bv<sup>r</sup>) quoted such an argument of his adversaries: “*Aiunt linguam horridam esse [...] nec multo certius eam percipi posse, quam si nunc velimus Aegyptiorum Hieroglyphicas literas interpretari.*” (Translation: They say that the language is rough [...] and that it cannot be understood much better than if we now would want to interpret the hieroglyphic letters of the Egyptians).

9. Cf. François (2003 & 2005).

its relation to other Semitic languages (especially Arabic, Aramaic, and Syriac) that were then more intensively studied. The character of these orations was less apologetic and more ceremonial. One of these later orations is the *Linguae Hebraicae encomium* (Louvain, 1614) by Andreas Valerius (1588-1655), professor of Hebrew at the Louvain Collegium Trilingue<sup>10</sup>.

The ideological program for the Collegium Trilingue was set by the famous humanist Desiderius Erasmus in his *Ratio seu compendium verae theologiae* (Basel, 1519). In this programmatic booklet, Erasmus defended the emphasis on language education as a gateway to a more philologically sound theology. The first professor of Hebrew, the Spanish converted Jew Matthaeus Adrianus (c.1475-after 1521), delivered an oration at Louvain in 1519 that was later published in Wittenberg, in which he defended the utility of the study of Hebrew against his opponents, who were clearly the Louvain theologians. Jacobus Latomus wrote his *De trium linguarum et studii theologici ratione dialogus* (Antwerp, 1519) in reaction to Adrianus' oration, but he also aimed his attack at the more influential *Ratio* of Erasmus. The Englishman Robert Wakefield (d.1537), the second professor of Hebrew at the Collegium Trilingue, taught only a few months at Louvain, but in 1524, he too delivered an oration on Hebrew, Arabic, and Chaldaic.

It is in this context that the anonymous oration in praise of the Hebrew language (*Oratio in laudem linguae Hebraicae*) that interests us was written. The autograph of the text, which does not mention its author's name, is preserved in a sammelband that was compiled by the sixteenth-century Hellenist Theodoricus Langius (d.1578). A later annotator added a title and listed some possible authors, but none of these names can be definitively tied to the oration. The author writes Hebrew with an elegant Sephardic cursive hand, which means he was relatively well-versed in the language<sup>11</sup>. He also directs his speech to an audience of students, as he refers to a class on the Book of Psalms, which suggests that he taught Hebrew at the Collegium Trilingue. Most remarkably, the author showed a profound interest in Christian

---

10. Studied in detail by Denecker & Van Hecke (2019). See also Zwiep (1993).

11. See Van Hecke & Feys (2017).

Cabbala<sup>12</sup>. In the sixteenth century, Christian Hebraists developed their own version of this reading strategy and used it to prove the prefiguration of Christ in the Old Testament. An extensive part of his brief appraisal of Hebrew is taken up by cabbalistic arguments aimed at proving that much of the original meaning is lost in translation. This interest in Christian Cabbala was particularly in vogue at the beginning of the sixteenth century and was famously exercised by Pico della Mirandola (1463-1494) in Italy, Johannes Reuchlin (1455-1522) in Germany, and Guillaume Postel (1510-1581) in France<sup>13</sup>. As for Louvain, there are few Hebraists who had an outspoken interest in Cabbala; one interesting example is Andreas Masius (1514-1573), who had studied at the Collegium Trilingue. In his notebook, which contained mainly self-written Hebrew poems, appears a list of cabbalistic works in Hebrew in which he was interested<sup>14</sup>, though Masius never involved these in his exegetic works.

### 3. “Lingua simplicissima”: simplicity as argument for learning Hebrew

The author begins his oration with a witty anecdote from the *Lingua* (Antwerp, 1525), a work by Erasmus, which itself is taken from Plutarch’s *Apophthegmata*. The anecdote relates that an orator was about to begin an oration on Hercules, when the Spartan king Antalcidas interrupted him by asking if there was anyone who could criticize Hercules. The message evoked here is that it would be absurd to praise someone or something which is already unanimously praised. He then proceeds with his main argument: describing the qualities of Hebrew. The first thing to remark about Hebrew is its divine nature: God created the language and taught it to Adam. This remark stems from a long debate begun in antiquity: is Hebrew the first language, invented by God, or one of the many languages that

---

12. Cabbala was a Jewish esoteric philosophy focused on looking for hidden meanings in Hebrew words and letters, which then would reveal a hidden reality. For a general introduction to Cabbala, see for instance Dan (2007) and Busi (1998).

13. For Postel’s interest in Christian Cabbala, see Secret (1964).

14. See Maleux (2020: 69).

appeared after the building of the tower of Babel? The church fathers, Augustine most notably, had no doubt that it was in fact the first language. Our author (like most humanists) holds the same opinion and declares that God created the language and taught it to Adam:

*Imo si penitius rem intueri placet, deprehendetis facile huius idiomatis ipsum Deum sanctum et benedictum autorem extitisse, quem legimus primum parentem nostrum Adam non alium quam hebraicum sermonem docuisse, qui eum simplicitatis ac castitatis admoneret: est enim haec lingua omnium linguarum simplicissima purissimaque, et ob hoc ipsum Deo quam gratissima, ac in explicandis divinae naturae mysteriis quam maxime conveniens atque idonea.*<sup>15</sup>

‘For if you would like to look at the matter more precisely, you will understand with ease that God Himself is the holy and blessed creator of this idiom; for we read that He taught none other language to our first ancestor Adam than Hebrew, and He made him aware of its simplicity and purity. For it is the most simple and pure of all languages, and therefore pleases most to God, and it is the most convenient and apt to explain the mysteries of the divine nature.’

God deemed this language suitable to teach mankind because of its *castitas* (purity) and *simplicitas* (simplicity). The qualifier *purissima* refers to the belief that Hebrew is a pure and true language without any defects (an idea known as the *hebraica veritas*, or Hebrew truth). Here, simplicity of language is used as illustration and proof of its divine nature and purity. The author nowhere provides grammatical examples of the simplicity of the Hebrew language, but takes it as a fact, as a consequence of its creation. There are, however, other cases in which grammatical examples have been used in defense of the language. In his oration on Hebrew, Arabic, and Chaldaic, Robert Wakefield (d.1537) claims that Hebrew is an easy language to learn, especially for Englishmen. He argues that the pronunciation of the language is very close to that of the English language:

*[A]d eius vocabula efferenda Gallis [...] Italis & Germanis multo nostrates sint aptiores. Quandoquidem ea aptitudine & facilitate modo operam dent,*

---

15. Ms. 744-55 p. 1.

*hebraica ad purum verba inspirant Angli ac de illorum ore resonat, qua media nati educatique Hierosolyma Iudei.*<sup>16</sup>

‘[I]n pronouncing these words, [...] our fellow countrymen are much better versed than the French, the Italians, and the Germans. Whenever they just pay attention to this aptitude and simplicity, the English breathe the Hebrew verbs purely, and it resonates from their mouth as if they were Jews born and brought up in the middle of Jerusalem.’

We also find an argument about the grammatical basis for Hebrew’s simplicity in college notes taken in the classes of Andreas Balenus (c.1484-1568), professor of Hebrew at the Trilingue from 1532 until 1568. As a rule, Hebrew words are formed by three consonants that carry the word’s basic meaning (the so-called *root*), while the vowels provide only additional information. Balenus put it this way:

*Quamquam in addiscendis cunctis fere linguis nihil sit aut difficilius aut utilius vocabulorum primitivorum exacta cognitione. In hæbreorum tamen lingua huius rei noticia nihil potest esse facilius, cum huius lingue primitiva vocabula sunt admodum pauca, atque adeo trium literarum dumtaxat omnia, quod sine grandi mysterio factum esse non putes.*<sup>17</sup>

‘Although in the learning of almost all languages nothing is more difficult or more useful than the exact knowledge of roots, in the Hebrew language nothing is easier than knowing this matter, since the roots of this language are very few and all consist of only three consonants, which you cannot imagine being made without great mystery. For as the Lord himself, the creator of this language, much rejoices in the threefold number of the persons, so too are all the roots of this language completed by the threefold number of the consonants.’

Here, too, we can see that the simplicity of Hebrew is related to its divine creator, suggesting Hebrew was some sort of divinely constructed language, designed to be easy and efficient. Balenus is far from the only Hebraist to make this observation. The Italian grammarian Santes Pagninus (1470-1541) ascribed the same Trinitarian perfection to the Hebrew root in his monumental grammar *Institutiones*

---

16. Wakefield (1524: 38). He furthermore notes in (transcribed) Hebrew that the French are not able to correctly pronounce the letter  $\psi$  ( $\xi$ ).

17. Ms. 8471-75 fol. 25’.

*Hebraicae* (Lyon, 1526)<sup>18</sup>. Balenus explicitly relates this divine feature to the concept of simplicity, which Pagninus, however, does not seem to suggest.

One of the clearest examples of basing Hebrew's simplicity on its grammar appears in the orations of the German Hebraist Hermannus Rennecherus (1550-after 1605), who taught Hebrew at Leiden University at the end of the sixteenth and the beginning of the seventeenth century. In 1603, he wrote an oration on the excellence of the Hebrew language (*oratio brevis in laudem hebraeae linguae*). In this oration, Rennecherus develops his argument in a well-structured manner, highlighting every aspect he considers characteristic of Hebrew, of which simplicity is an important one (*antiquitas, dignitas, sanctitas, castitas, utilitas, necessitas, brevitatis, facilitas, jucunditas*). He explains that its simplicity relates to its brevity<sup>19</sup> and enumerates many grammatical (morphological) features (like gender and number in nouns and the verb tenses, but also Hebrew accents) which are simpler in Hebrew than Latin and Greek, languages Rennecherus characterizes as being very difficult. If there are exceptions in Hebrew, he remarks that they are so few that they pose no difficulty to the learner. Moreover, he not only relates simplicity of language to brevity and conciseness, but also to aesthetic qualities.

#### 4. “Lingua purissima”: simplicity in conjunction with the perfect divine nature of Hebrew

When the Latin church fathers described Hebrew, they generally acknowledged it as a holy tongue and the oldest language, but that did not prevent them from calling it ugly and difficult, as we already

---

18. See Kessler-Mesguich (2013: 22-23 & 129-130).

19. Rennecherus 1603: 22: “*Facilitas huius linguae aliqua ex parte cum eius brevitate est coniuncta. Primo igitur in eo haec lingua est facilis, quod tam multa vocabula & voces non habeat quam graeca & Latina lingua, ut ante in eius brevitate dictum est.*” (Translation: This language's simplicity is partly related to its brevity. First of all this language is easy, because it does not contain so many words and expressions as Greek and Latin, like I remarked before on account of its brevity).

saw in the case of St. Jerome. Humanists were remarkably more positive, because for them the qualities of Hebrew were all related to each other. They theorized that it was the holiest of languages, since its creator was God, who taught it to Adam. Therefore, it was also by far the oldest language to have existed, as well as the most perfect language, since it was designed to be as such.

A perfect language for humanists was a language which was highly logical and consisted of short words reflecting reality<sup>20</sup>. This approach, known as iconicity, is articulated in Plato's *Cratylus*, where the eponymous character asserts the intrinsic link between the referred object and the word. Plato, however, did not seem to share this view and propagated through Socrates a more moderate view that most words do not share this link with reality, though there are unmistakably words that do<sup>21</sup>. In the Renaissance, humanists chose to follow Cratylus' approach to forming etymologies, combined with Varro's etymological procedures of shifting letters. This led certain humanists to develop etymologies based on Hebrew roots, which to our ears sound quite outlandish<sup>22</sup>. The quality of simplicity is therefore a consequence of the aforementioned qualities.

The idea that perfection can be combined with simplicity is clearly formulated in the following passage from the anonymous Louvain oration:

*Quod si non nobilitate originis commendari videbitur, certe hoc inficiari non possunt, quin propter sacram scripturam, quam hebraice scriptam esse, apud omnes in confesso est, permultum laudis commendationisque accipiat. Si enim propter rerum cognitionem ut bonarum disciplinarum, artiumque scientiam, putamus grecam linguam aut latinam esse utilem, idem de lingua sancta cur non audebimus dicere, quae perfecta est et absoluta neque ullo modo potest fallere, sed omnia simpliciter ac plane ob oculos ponit, quae ad intelligentiam scripturarum sunt necessaria.*<sup>23</sup>

---

20. A prominent defender of this theory was the Dutch humanist Johannes Goropius Becanus, though he considered the Antwerpian dialect instead of Hebrew to be the oldest language on these same grounds. See Deneire & Van Hal (2006: 24).

21. See e.g., Law (2003: 20-23).

22. For a seventeenth-century example of this practice, see Eco (1993: 92).

23. Ms. 744-55 p. 3.

‘If it does not seem to be praiseworthy because of its noble origin, then certainly they cannot deny that it receives all praise and honor because of the Holy Scripture, which as everyone knows is written in Hebrew. For if we think that Greek and Latin are useful for both practical knowledge and intelligence in the good disciplines and arts, why do we not dare to say the same about the holy language, which is perfect and absolute and cannot be wrong whatsoever, but poses simply and clearly before one’s eyes everything that is necessary to understand scripture.’

Because Hebrew is a perfect language without flaws, it is not ambiguous and therefore clear and simple in its idiom. The author proceeds to give examples of ambiguities in the Greek translation of the Biblical Hebrew text that had caused multiple errors in the Latin translation.

Yet, the author primarily emphasizes cabbalistic arguments. In his view, the most important reason to learn Hebrew is to be able to read the Old Testament without reliance on translations, since a translation cannot possibly transfer every nuance as precisely as the original text. In addition, the Hebrew text possesses not only meaning on the superficial level, but also hidden meanings behind the Hebrew letters, a theory known as Cabbala. Our author offers a couple of examples to prove his point: the first word of the Hebrew bible, בראשית (*bərē’sīt*, “in the beginning”) also contains the words בר אשית (*bar ’āšīt*), which means “I will bring forth a son”<sup>24</sup>. This hidden meaning is obviously impossible to perceive via a translation; therefore, one should learn Hebrew to become aware of the cabbalistic interpretations. The use of Cabbala to recommend the usefulness of the study of Hebrew is reminiscent of the words of the famous Christian Hebraist Johannes Reuchlin, author of the *De arte cabalis-*

---

24. Ms. 744-55 p. 8: “*In prima si quidem dictione, Deus pater stato certoque tempore, se daturum nobis filium promittit. Nam vocabulum בְּרֵאשִׁית quod proprie, in principio significat, sic prolatum בְּרֵאשִׁית significat filium dabo: בר bar enim filius est, unde et Petrus Simon bariona, hoc est, filius columbae, a Christo appellatus est אשית autem dabo significat.*” (Translation: Indeed, in the first word, God the Father promises to give us a son on a fixed and certain moment. For the word *bərē’sīt* – which in itself means “in the beginning” – means “I will give a son” when pronounced as *bar ’āšīt*; for *bar* means “son,” whence Peter Simon is also called *Bariona*, which means “son of the dove;” *’āšīt* on the other hand means “I will give”).

*tica* (Hagenau, 1517): “[the cabbalistic hidden meanings] cannot be successfully discerned if not by a scholar of Hebrew”<sup>25</sup>. The author’s focus on Cabbala should be understood in the light of Reuchlin’s book, which was probably read in Louvain<sup>26</sup>. These untranslatable hidden meanings would most probably have stirred the students’ imagination, and it moreover strengthens the idea of the divine nature of the language: God himself laid these hidden meanings in the Hebrew letters, and therefore no other language more purely expresses the divine message.

## 5. Conclusion

In conclusion, Hebrew’s apologists do not describe the language as “simple” to lure students. The concept of simplicity is closely related to the concept of the perfect language, which Hebrew was thought to be during the early age of humanism. The holy and pure nature of Hebrew implied that its structure was highly logical, and thus it should be easy to understand. At the same time, behind every Hebrew letter there lurked other meanings that could be interpreted only with a sound knowledge of Christian Cabbala. Only later, towards the beginning of the seventeenth century, did scholars begin to doubt the traditional primeval position of Hebrew<sup>27</sup>. The idea of Hebrew as a simple and universal language, however, continued to exercise influence until the nineteenth century, when the quest for the perfect language was peaking (resulting in the creation of artificial languages like Volapük and Esperanto) and when Modern Hebrew was re-invented as a national language for the Jewish people<sup>28</sup>.

---

25. “*a nemine prorsus intellecta nisi hebraice praedocto.*” Reuchlin (1506: 4).

26. In the 1543 inventory of a bookshop located near the Trilingue, the *De arte cabbalistica* is included. See Delsaerd (2001).

27. See Van Hal (2010).

28. On this particular history, see Halperin (2012).

## References

### Manuscript sources

- Anonymous. *Oratio in laudem linguae Hebraicae*. Royal Library of Belgium: mss. 744-55.
- Anonymous. *Annotationes quaedam ad hebraeae linguae noticiam*. Royal Library of Belgium: mss. 8471-75.

### Primary sources

- Adrianus, Matthaeus. 1520. *Oratio quam Louanii habuit de linguarum laude*. Wittenberg: per Ioh. Grunenberg.
- Andreas, Valerius. 1614. *Collegii Trilinguis Buslidiani, in Academia Lovaniensi, exordia ac progressus. Et linguae Hebraicae encomium*. Lovanii: Typis Philippi Dormalii.
- Clenardus, Nicolaus. 1529. *ליה הדיקדיק* Tabula in grammaticen Hebraeam. [Louvain]: Theodo[ricus] Martin[us].
- Erasmus, Desiderius. 1519. *Ratio seu compendium verae theologiae*. Apud inclytam Basileam.
- Erasmus, Desiderius. 1525. *Lingua per Des[iderium] Erasmum Roterodamum, opus nouum, et hisce temporibus aptissimum*. Antuerpiae: apud Michaellem Hillenium.
- Latomus, Jacobus. 1519. *De trium linguarum et studii theologici ratione dialogus*. Antuerpiae: per Michaellem Hilleniu[m].
- Levita, Johannes Isaac. 1559. *Defensio veritatis Hebraicae sacrarum scripturarum, aduersus Libros tres Reueren[di] D. Vvilhelmi Lindani S.T. Doctoris, quos de optimo Scripturas interpretandi genere inscripsit*. Coloniae: apud Iacobum Soterem.
- Mosellanus, Petrus. 1518. *De variarum linguarum cognitione paranda oratio*. Lipsiae: in officina valentini Schvmann.
- Pagninus, Sanctes. 1526. *Hebraicae Institutiones [...]*. Lyon: Antoine du Ry.
- Rennecherus, Hermannus. 1603. *Oratio breuis et succincta in laudem sanctae et reuerendae Hebraeae linguae*. Lugduni Batavorum: apud Thomam Basson.
- Reuchlin, Johannes. 1506. *De rudimentis hebraicis libri tres*. [Pforzheim]: [Thomas Anshelm].
- Reuchlin, Johannes. 1517. *Ioannis Reuchlin Phorcensis LL. Doc. De arte cabalistica libri tres Leoni X. dicati*. Hagenau: apud Thomam Anshelmum.
- Wakefield, Robert. 1524. *Oratio de laudibus & utilitate triu[m] linguar[um] Arabice Chaldaicae & Hebraice atq[ue] idiomatibus hebraicis quae in utroq[ue] testame[n]to i[n]ueniuntur*. London: apud VVinandum de Vorde.

- Winshemius, Vitus. 1549. *Orationes duae de stvdiis linguae Ebraicae*. Vitebergae.
- Wright, Frederic A. 1933. *Jerome, Select letters*. Cambridge: Harvard university Press.

## Secondary sources

- Barr, James. 1967. St. Jerome's appreciation of Hebrew. *Bulletin of the John Rylands University Library of Manchester* 49. 281-302.
- Busi, Giulio. 1998. *La Qabbalah*. Roma-Bari: Editori Laterza.
- Dahan, Gilbert. 1990. *Les intellectuels chrétiens et les juifs au Moyen Âge*. Paris: Editions du Cerf.
- Dan, Joseph. 2007. *Kabbalah: a very short introduction*. Oxford: Oxford University Press.
- Delsaerd, Pierre, 2001. *Suam quisque bibliothecam. Boekhandel en particulier boekenbezit aan de oude Leuvense universiteit 16de tot 18de eeuw*. Leuven: Universitaire Pers Leuven.
- Denecker, Tim. 2015. Language attitudes and social connotations in Jerome and Sidonius Apollinaris. *Vigiliae Christianae* 69. 393-421.
- Denecker, Tim & Pierre Van Hecke. 2019. Why learn Hebrew? Text and translation, with an introduction and commentary, of Valerius Andreas' *Linguae Hebraicae Encomium* (1614). *Lias* 45. 45-111.
- Deneire, Tom & Toon Van Hal. 2006. *Lipsius tegen Becanus*. Amersfoort: Florivallis.
- De Vocht, Henry. 1951-55. *History of the Foundation and the Rise of the Collegium Trilingue Lovaniense 1517-1550* (4 vols.). Louvain: Publications Universitaires.
- Dunkelgrün, Theodor. 2017. De bekeerling als onmisbare schakel tussen joodse en humanistische geleerdheid. *Erasmus' droom: Het Leuvense Collegium Trilingue 1517-2017. Catalogus bij de tentoonstelling in de Leuvense Universiteitsbibliotheek, 18 oktober 2017 – 18 januari 2018*, ed. by Jan Papy. Leuven: Peeters. 416-419.
- Eco, Umberto. 1993. *La ricerca della lingua perfetta nella cultura europea*. Roma – Bari: Laterza.
- François, Wim. 2003. The plea by the humanist Petrus Mosellanus for a knowledge of the three biblical languages. A Louvain perspective. *Revue d'Histoire Ecclésiastique* 98. 438-481.
- François, Wim. 2005. 'Ad divinarum rerum cognitionem'. Petrus Mosellanus and Jacobus Latomus on biblical or scholastic theology. *Renaissance and Reformation* 29. 13-47.

- Halperin, Liora R. 2012. Modern Hebrew, Esperanto, and the quest for a universal language. *Jewish Social Studies* 19. 1-33.
- Kessler-Mesguich, Sophie. 2013. *Les études hébraïques en France de François Tissard à Richard Simon (1508-1680)*. Geneva: Droz.
- Law, Vivien. 2003. *The History of Linguistics in Europe: From Plato to 1600*. Cambridge: Cambridge University Press.
- Maleux, Maxime. 2020. On Hebrew nymphs and aqueducts. Two Hebrew poems by Andreas Masius. *Lias: Journal of Early Modern Intellectual Culture and Its Sources* 47. 67-87.
- Papy, Jan, ed. 2017. *Erasmus' droom: Het Leuvense Collegium Trilingue 1517–2017. Catalogus bij de tentoonstelling in de Leuvense Universiteitsbibliotheek, 18 oktober 2017 – 18 januari 2018*. Leuven: Peeters.
- Secret, François. 1964. *Les kabbalistes chrétiens de la Renaissance*. Paris: Dunod.
- Van Hal, Toon. 2010. “Moedertalen en taalmoeders”. *Het vroegmoderne taalvergelijkende onderzoek in de Lage Landen*. Brussel: Koninklijke Vlaamse Academie voor Wetenschappen en Kunsten.
- Van Hecke, Pierre. 2018. Omnium linguarum purissima: The study of Hebrew at the Collegium Trilingue. *The Leuven Collegium Trilingue 1517-1797. Erasmus, Humanist Educational Practice and the New Language Institute Latin-Greek-Hebrew*, ed. by Jan Papy. Louvain: Peeters. 155-182.
- Van Hecke, Pierre & Xander Feys. 2017. Anonieme lofzang op de Hebreeuwse taal. *Erasmus' droom: Het Leuvense Collegium Trilingue 1517–2017. Catalogus bij de tentoonstelling in de Leuvense Universiteitsbibliotheek, 18 oktober 2017 – 18 januari 2018*, ed. by Jan Papy. Leuven: Peeters. 419-427.
- Zwiep, Irene E. 1993. The unbearable lightness of Hebrew. Didactic trends in seventeenth-century Dutch grammars of Biblical Hebrew. *Five hundred years of foreign language teaching in the Netherlands 1450-1950*, ed. by Noordegraaf, Jan & Frans Vonk. Amsterdam: Stichting Neerlandistiek VU. 27-55.

# Representations of linguistic simplicity in prehistoric fiction

**Mat Pires**

Université Bourgogne–Franche-Comté, Besançon, CRIT EA3224

---

## ABSTRACT

This chapter investigates the issue of linguistic simplicity as it appears in prehistoric fiction, and specifically the direct speech incorporated into such fiction. It first discusses the different approaches to foreign-language dialogue in the novel, contrasting *transfer* and *translation*. In most cases of prehistoric fiction, translation from the imaginary source language is adopted. The features of this “translated” speech are analyzed with regard to their complexity or simplicity compared to standard idiom. The following such simplifications are presented and assessed: avoidance of personal deixis; use of reconstructed descriptive lexical items characterized by transparent meaning; morphology of transferred names of people; avoidance of politeness features, in the form of sentence-words and deontic modalization. Some verbal features are compared with similar features observed by Daniel Everett (2009) in the Pirahã language.

## KEYWORDS

prehistoric fiction, linguistic simplicity, comic dialogue, novelistic dialogue, oral and written language, imaginary language

## RÉSUMÉ

Ce chapitre aborde la question de la simplicité linguistique à travers l'exemple de la fiction préhistorique, analysant plus particulièrement les répliques de discours direct qui figurent dans cette fiction. Il présente d'abord deux manières possibles de traiter ce dialogue étranger, à savoir le *transfert* et la *traduction*. Dans la plupart des cas le dialogue préhistorique prend la forme d'une traduction depuis une langue-source imaginaire. Les spécificités de ce discours « traduit » sont analysées pour identifier les traits

de simplicité ou de complexité qui pourraient le différencier de la langue standard contemporaine. Nous discutons ainsi l'évitement de la déixis personnelle ; la présence de substantifs « descriptifs » à sémantisme transparent ; la forme des anthroponymes transférés ; et l'absence de marques de politesse sous forme de mots-phrase ou de modalisation déontique. Certains traits verbaux sont rapprochés de traits observés dans la langue Pirahã, dans les travaux de Daniel Everett (2009).

#### MOTS-CLÉS

fiction préhistorique, simplicité linguistique, dialogue de la bande dessinée, dialogue romanesque, langue orale et langue écrite, langue imaginaire

---

The approach to simplicity in language in this article has the unusual feature of not being concerned with actual data collected in the field for a given language, nor even with metalinguistic descriptions of such data. Instead it looks at *representations* of unknown and indeed unknowable languages, seen through the prism of contemporary idioms into which they are *understood* to be *translated*. These representations occur in works of literature, which itself represents a further filter in terms of the written channel and the specific genre. With such a praxis whether one is dealing with *a language* becomes a moot point. However, as we will see, there is undeniably a quest for linguistic simplicity, which manifests itself in multiple linguistic features.

The genre in question is the *prehistoric novel*. It is defined as “a genre having some common ground with science fiction, with the specificity of enacting events taking place in Prehistory, with varying degrees of realism”. It certainly appears as a well-represented, popular type of fiction: the website *Les Romans de la Préhistoire*, which provided that definition, currently lists 420 novels and plays, and 352 comic titles. The earliest work dates back to 1836, but according to the compiler the genre really took off in the 1860s. It is dominated by two language-cultures: French and English (the same observation is made in Ruddick's (2009: 207) monograph on the subject), for reasons at least partly to do with the celebrity of prehistoric sites, both long-established ones such as Stonehenge, and more recent ones, such as the painted caves uncovered between the later nineteenth century and the discovery of Lascaux in 1940.

Among the most emblematic titles in the genre are J.-H. Rosny aîné's *La guerre du feu* (1911), William Golding's tale of the encounter between Neanderthals and modern humans in *The Inheritors* (1955), the comic series *Rahan*, by Leheureux and Chéret (first published 1969), and Roy Lewis's pastiche – a clear signal of the existence of a genre – *How I Ate My Father*, from 1960.

Our analysis concerns the forms of direct speech turns in these works<sup>1</sup>. Direct speech poses two challenges to any author with regard to realism. The first, which concerns any novel incorporating speech, is the *written orality paradox*: the speech of fictional characters is considerably scripturalized, lacking the in-progress, inchoate form of natural spontaneous speech. The second is the *translation paradox*, which concerns any novel set in communities which do not use the language in which the novel itself is written (henceforth “the writing language”). The author needs an understanding with the reader that the writing language is to be taken as representing the language spoken in the setting (henceforth “the setting language”).

Writers often pay lip-service to the translation paradox by incorporating terms from the setting language, a practice caricatured by Paul Theroux in one of his travelogues:

In Spanish I said, “It is quiet here.”

That was the first time on my trip that I spoke Spanish. After this, nearly every conversation I had was in Spanish. But in the course of this narrative I shall try to avoid affecting Spanish words, and will translate all conversations into English. I have no patience with sentences that go, “‘*Carramba!*’ said the *campesino*, eating his *empanada* at the *estancia*...” (Theroux 1980: 49, italics his).

However, Cormac McCarthy, again with English and Spanish, leaves all direct speech in the setting language, either choosing sentences which are likely to be familiar to his target readership, or else using the narrative to elucidate:

He squatted in front of them.

Quiero comprar una trucha, he said.

---

1. We have preferred to limit the analysis to turns of dialog in order to center the discussion around language, though narrative passages may imply speakers' recognition of certain features or cognitive processes.

They nodded. The one named Faustino spoke.  
 Cúanto [*sic*] dinero tienes?  
 Cuarenta y cinco pesos.  
 The indian came over.  
 Siéntate, he said.  
 The indian sat. (McCarthy 1993: 196-97)

The written orality paradox and the translation paradox are compounded in prehistoric fiction – in fact there may even be a further *language paradox* since the earliest concrete traces suggesting the existence of language – those of proto-writing – go no further back than ten millennia (Li *et al.* 2003; Tallerman & Gibson 2013). Existence of language prior to this is *by definition* based on extrapolations from physiological features or genes, deductions from data on the evolution of hominids and evidence of their lifestyle, technological capabilities, cultural traces and so on (DePaolo 2000: 431).

These contrasting narrative postures, taken from novels set in contemporary periods, suggest the range of options open to the prehistoric novelist.

1. *Transfer*: reproduce (invented) setting language (+ realistic; – comprehensible) (McCarthy option)
2. *Translation* into the writing language...
  - 2a ... adopting a standard idiom (– realistic; + comprehensible) (Theroux option)
  - 2b. ... incorporating modifications to create a sense of otherness (~ realistic; ~ coherent)

We find representatives of each in the prehistoric novel.

## 1. Transfer

A simple example of transfer is found in Pierre Pelot's *Sous le vent du monde*, a series of novels advertised as having benefited from the “*collaboration scientifique*” of the French paleoanthropologist Yves Coppens, and set in periods ranging from 1.7 million to 32,000 years ago. The novels use transfer systematically, as for example in this extract from a story narrated by Neh-Ishri'n':

Les mots tombaient hors de sa bouche comme une grande roucoulade.

– *Ataniktô-nik-nik ash hr'iaw draloa. Nakoa-xri' Nakoa-atanik-edri. Iw'oa-hua-hua hr. Iw'oa-hua-hua Oadwi han, nar, ra-nar'h, kiow tô-nik-ok. V'rhx nakoa-atanik-edri-neh! Iw'oa-hua-hua Oadwi nakoa-xir' ra-nar'h.*

Ils l'avaient laissée raconter et ils approuvaient en balançant la tête, les dents largement découvertes. (Pelot 1997: 119)

The problem of incomprehension is alleviated textually: a character has recently discovered a dead animal, a *nakao-xri'*, and with similar occurrences we learn that *âtanik* means “lots”, and that *v'rhx* means “leave” or “flee”. So, the reader picks up some of the lexicon, and finally perceives a certain complexity - for example making out instances of morphological marking.

## 2. Translation

Unlike Pelot, most writers seem to adhere to the notion that prehistoric language was simpler than contemporary language. Ruddick (2009: 173) suggests that the idea of simplicity originates in Victorian times, when the notion of a rudimentary prehistoric language bolstered the idea of the innate superiority of present-day humans, whose privileged position had been undermined by the theory of Evolution. This is posed or theorized in many places, for example by Jack London:

Perhaps, all told, the Folk in that day had a vocabulary of thirty or forty sounds. I call them SOUNDS, rather than WORDS, because sounds they were primarily. They had no fixed values, to be altered by adjectives and adverbs. These latter were tools of speech not yet invented. Instead of qualifying nouns or verbs by the use of adjectives and adverbs, we qualified sounds by intonation, by changes in quantity and pitch, by retarding and by accelerating. The length of time employed in the utterance of a particular sound shaded its meaning. We had no conjugation. One judged the tense by the context. (London 1906-07, chap. 4)

or in Roy Lewis's spoof:

[Endogamy] means we don't extend our very small vocabulary and our limited grammar; which in turn means a restricted power of abstraction. Language precedes and breeds thought, you know; and it is really little more than a courtesy to call a language the few hundred substantives we possess, the score of all-purpose verbs, the poverty of prepositions

and postpositions, the continued reliance upon emphasis, gesture and onomatopoeia to eke out shortages of cases and tenses. (Lewis 1993: 105)

Despite such theorizing, direct speech in prehistoric fiction is frequently standardly complex, and almost always in the writing language. This creates a paradox since it stands for a supposedly *simple* proto-language. However, *translation* enables the contradiction to be resolved: the simpler language has *in translation* become more complex than its source. The flimsiness of this approach is pointed up by Lewis, whose characters' supposedly protean language manages to include among its few hundred nouns items like *postposition* and *onomatopoeia*. One character, "Uncle Ian", even mocks a tribe he dismisses as "Leftovers from the Miocene" with the pidgin-type phrase "Dat ape im belong big big spear" (*ibid.*: 94). So, Harris sees the issues raised by iconic portrayals of prehistoric language, and satirizes the "face saving" provided by the fig-leaf of translation<sup>2</sup>. This first approach, then, we could follow translation theorists and call *target-language biased*.

At the other end of the spectrum, we have *source-language biased* translation, where direct speech *is* in the writing language but is made unnatural in various ways so as to capture the supposed forms of prehistoric language. There are several ways this may occur.

### 3. Avoidance of personal deixis

The first is when deictic pronouns are replaced with referential third-persons, either for the speaker or the interlocutor. Such a trait features in the earliest prehistoric fiction: Ruddick, referring to Élie Berthet's *Romans préhistoriques* (1876), notes how one character, Red, "can speak, but like most dimwitted cavemen thereafter, [...] uses third person both to address an interlocutor ('Deer shall eat') and to refer to himself ('Red was strongest')" (2009: 175). This constructs personal deixis as a recent complexification, and we find its avoidance in many examples of otherwise target-language biased texts. In *La guerre du feu*, it is *almost* systematic – there are in fact eleven tokens of *je*,

---

2. The conflict does lead most authors tending toward standard idiom to make some "iconic" allowances, or otherwise attenuate the clash. Jack London, for instance, avoids direct speech, while Rosny provides no theory of language.

twelve of *tu* and 20 of *vous* – but these exceptional deictics seem to be motivated by a desire to avoid repetition or awkwardness. For example, in:

Si les mammoths quittent le pâturage, fit Nam, *je* réveillerai le fils du Léopard. (Rosny 1911, 2.7)

*Je* suis Naoh, fils du Léopard. (*ibid.*, 3.9)

the *je* tokens avoid repetition (“fit Nam, Nam réveillera...”; “Naoh est Naoh”). Elsewhere *tu* is used in addressing a mammoth:

– [...] nous avons arraché ces plantes pour que tu puisses en faire ta pâture. Ainsi tu sauras que les Oulhamr sont les amis du mammoth. (“tu” = mammoth, *ibid.*, 2.3)

Such simplification runs from the 19<sup>th</sup> century to the present, as this extract from Marc Klapczynski’s *Aô l’homme ancien* (2003) shows. Âki-naâ is addressing the hunter Ka-Maï, member of her tribe:

– Âki-naâ va chasser avec Aô.

Cette seule phrase suffit pourtant à déclencher le courroux du vieux chasseur qui maugrée :

– Âki-naâ sait qu’elle ne doit pas tuer d’animaux. (Klapczynski 2003: 155)

and it is omnipresent in the comic book tradition exemplified by *Rahan*. Golding’s *Inheritors* (1955) is notable in *not* adopting this simplification<sup>3</sup>.

#### 4. Immediacy of experience

Since most prehistoric fiction dialog is notionally *translated* from a source language, it is in general terms no less complex than standard English or French. William Golding’s *The Inheritors*, however, does use a perceptible “adaptation” of the spoken idiom. Ruddick (2009: 178) goes so far as to describe the work as staging an ‘encounter between transitional speech and articulate language,’ the former being that of the Neanderthals and the latter that of the modern humans whose viewpoint is adopted in the final chapter. The reader is alerted to the

---

3. For a persuasive, sociopragmatic approach to Golding’s characters’ language, which sees it as not simplistic at all but differently adapted to its environment, see Mandala 2018.

language issue in various ways: the 17 occurrences of the phrase “I have a picture” which seek to capture the Neanderthals’ cognitive processes, or the discovery of metaphor in the passage beginning “Lok discovered ‘Like’ ” (Golding 1955: 194-195).

How might one measure such simplicity? Since it is difficult to establish a comparable corpus of novelistic speech against which to gauge possible deviance, we have used as a point of comparison a language noted for its many simple characteristics: the Pirahã language of the Amazon. The work of Daniel Everett has brought to light a series of constraints proper to this idiom, including the requirement that utterances be assertive, the exclusion of recursion of all types (relative clauses, conjunction and disjunction), and the absence of numerals or color adjectives. Verbal forms which locate processes outside the moment of utterance (exemplified in English by *I had eaten*, or *I will have eaten*) are not found (2009: 132-33), nor is a passive voice (*ibid.*: 223). Everett equates these features with what he calls the *Immediacy of Experience Principle* (*ibid.*: 115 ff.): “Declarative Pirahã utterances contain only assertions related directly to the moment of speech, either experienced by the speaker or witnessed by someone alive during the lifetime of the speaker” (*ibid.*: 132).

Certain of these features are indubitably present in the dialog of *The Inheritors*. Complex, recursive syntax, for one, is readily summoned in both its communities: “I have a picture of the time when honey will run out of this crack in the rock so that you can taste honey off your fingers—so!” (Neanderthal, in Golding 1955: 51); “You and he gave my child to the devils and they have given me back a changeling who does not see or speak” (human, *ibid.*: 229). However, several other features do appear to give rise to avoidance, albeit with occasional exceptions. Counting occurs in a single speech turn (“One day. Perhaps two days. Not three”, *ibid.*: 13), while color terms are limited to three tokens describing hyenas as “the yellow ones” (*ibid.*: 53, 55, 119). Non-present perfective verbal forms occur only twice (“And armfuls of meat that a cat had killed”; “If we had not we should have died”, *ibid.*: 59, 228), and there are only two instances of passive voice (“I have a picture of what is to be done”, “That is a new thing. But it was spoken”, *ibid.*: 36, 47).

However, the position defended by Ruddick that simplicity and complexity are articulated through an opposition between the language of the two species is not borne out by the dialog in the novel. The dialog of the sapiens characters is found in just 22 speech turns in the 12th and last chapter of the novel, a limited source of material. The only complexity unique to this sub-corpus is the combination of modal verb and past tense, which occurs as its first and third turns: “What else could we have done?” and “If we had not we should have died” (*ibid.*: 227, 228). In these utterances, multiple values complexify the bare verbal notions of *do* and *die*: past tense, interrogation, hypothesis, deontic and epistemic modalization, perfectivity. It is conceivable that the semantic depth of these turns serves to herald the arrival of the narrative among the modern humans.

## 5. The descriptive referent

Despite the translation principle, certain referents found in the prehistoric environment are named descriptively, in a way which brings out the semantic structure of the prehistoric-language name:

<i>eagle:</i>	<i>bec-crochu</i>
<i>antelope:</i>	<i>deux-cornes, fines-cornes, longues-cornes</i>
<i>afterlife:</i>	<i>territoire des ombres</i>
<i>autumn:</i>	<i>saison-des-feuilles-jaunes.</i>

These examples are taken from the comic book *Rahan*, where such naming is a staple – the website [rahan.org](http://rahan.org) lists over seventy examples. Footnotes are even provided in some cases (Figures 1, 2).



Figures 1 (left) and 2 (right):  
Examples of footnotes from *Rahan* (n° 24, p. 32, 11).

Occasional examples do occur outside *Rahan*: “la bête-aux-dents-déchirantes” (“lion”, *La guerre du feu*, 1.4), or “the yellow ones” mentioned above. Most such lexical items are animals, with a few tools, foods, natural features or elements of cosmology. Nouns dominate: the list compiled on *rahan.org* contains only two verbs: *nager* (“ramper sur l’eau”) and *tituber* (“avoir les jambes creuses”).

In these examples simplicity is articulated by the symbolic re-statement of a semantic transparency understood to have been lost to language change over time. Transparent terms for natural phenomena do of course exist in modern idiom (such terms for fauna include *fly*, *blackbird*, *woodpecker*), but the vast majority of lexical items are not descriptively transparent in this way. Language users are however aware that etymology can provide a way back to such clarity. Evoking the existence of such source forms through their literal translation suggests both an edenic simplicity, and the antiquity of the language.

## 6. Anthroponyms

The *transfer* (unmodified reproduction) of character names is a standard feature of almost all fiction translation<sup>4</sup>. In prehistoric settings, these items are generally the sole iconic representations of prehistoric language, making them an important authenticating counterweight to the modern idiom in which the narrative and even the dialog are couched. A selection of such anthroponyms is given in Table 1.

<b>I-ka-eh</b>	Arcelin, <i>Solutré</i> (1872)
<i>Naoh, Nam, Goûm, Gaw, Faouhm</i> <b>Gammla</b>	Rosny aîné, <i>La guerre du feu</i> (1911)

4. See Ballard (2001: 18). The rare exceptions concern unsituated characters such as those of certain fairy-tales (*Peter and the Wolf*, *Pierre et le loup*, etc.), or famous names which have been naturalized in other languages (Michelangelo, *Michel-Ange*).

<i>Abaho, Abhân, Cush, Dhân, Eko, Gômer, Gotch, Khôm, Noûm, Rhâ, Sick, Tâni, Thor, Touati, Uri</i> <b>Enéa, Mâm, Oûma, Rama, Saga, Tânit, Tsilla, Zaïra</b>	Cénac, <i>Cavernes de la rivière rouge</i> (1967)
<i>Dahok, Êrak, Ghoho, Oonou, Pahouk, Powang, Rahan, Vràah, Xoar, Yako, Zumka</i>	Lécureux, <i>Rahan</i> , 24 (1977)
<i>Aô, Aki-naâ, Ka-mai, Ma-wâmi ; Owk, Bao, Taoum, Ayawakoum, Naâm, Zintaka, Awaki, Yâgwa, Napatchi, Wampa</i> <b>Iktia, Aïki, Waâ, Kima, Pakili</b>	Klapczynski, <i>L'odyssée du dernier Neandertal</i> , (2003-2010, v. 1 ; 2)
<b>Nî-éi, Moh'hr ; Oohki'a, Boohr'am, Oughuaq ; Èheni ; Dohuka, Naobah, Aruaeh, Tuhi-Horea</b>	Pelot, <i>Sous le vent du monde</i> (1996-2001, v. 1 ; 3 ; 4 ; 5 ; selection only)
Neanderthal: <i>Lok, Fa, Liku, Mal, Ha, Nil.</i> Sapiens: <i>Tuami, Vivani, Twal, Marlan, Tanakil, Vakiti</i>	Golding, <i>The Inheritors</i> (1955)
<i>Dorv, Zoug, Grod, Brun, Vorn, Broud, Goov, Creb, Mogur</i> <b>Iza, Ayla, Uka, Ebra, Oga, Ovra</b>	Auel, <i>Clan of the Cave Bear</i> (1980)

Table 1: Transferred anthroponyms in prehistoric fiction (male in italics, female in bold).

Overall these items present a mixture of complexity and simplicity. Some are basic – *Grod, Yako, Abaho, Bao, Êrak* – and feature the simple letter-sound correspondence typical of transliteration. Digraphs, however, are of uncertain value: is *Zoug* /zəʊg/ or /zavɡ/, or even /zʌg/ (like *Doug*)? In French novels they may signal an English-based standardized transcription: /ʃ/ in *Cush* or /u/ in *Oohki'a* and *Boohr'am*, although the opposition in *Oonou* remains puzzling. Non-alphabetical elements are even more resistant to interpretation. The circumflex, found in all but one of the French novels, cannot be phonetic in <î> or <û>; one can speculate that the markedly diachronic value of this ‘memorial accent’, as Ferdinand Brunot (quoted Cerquiglini 1995) termed it, is being drawn on. Pelot’s use of apostrophe does appear phonetic (announcing a glottal stop for example), though the name

*Moh'hr* resists this interpretation. The hyphens followed by lower case initials which appear in many names suggest an underlying semantic value (the properhood of the lower case element appearing uncertain).

One clear example of motivated complexification is found in *The Inheritors*. Here the sapiens' superior sophistication over the Neanderthals – a central theme of the novel – is reflected in their names, which are longer both in literal and syllabic terms.

*Translation* is only available where the name has semantic value. Examples are *Patte de Tigre*, one of the two characters in Arcelin's *Solutré* (curiously, the other has a complex transferred name, *I-ka-eh*), and "Cheveux-de-feu" which an online lexicon indicates as a translation of *Rahan* (Rahan.org). There are even combinations of the two: "Aghoo-le-velu" and "Goûn-aux-os-secs" (Rosny), "Saga-la-blonde" (*sic*) in Cénac. A similar transparency is served through filial naming, whether literally (Rahan as "fils de Crao") or metaphorically (*La guerre du feu*'s Naoh, Aghoo, Nam, Gaw being respectively "fils du Léopard", "fils de l'Aurochs", "fils du Peuplier" and "fils du Saïga").

Overall then, the only onomastic choice is between transfer and description. Substituted contemporary names would clearly disrupt the diegesis, a point demonstrated parodically in the skit *How I Ate My Father*, where the fourteen-year-old narrator is named Ernest<sup>5</sup>, and his tribal compatriots include Elsie, Oswald, Griselda, Uncle Ian and even Uncle Vanya.

Though I have dwelt on individual anthroponyms, collective appellations are similarly divided between transfer and description. On one hand semantic transparency is served by the Clan of the Cave Bear (Auel), Clan des Hommes du lac (Klapczynski), or Clan du Mont Bleu (Lécureux); on the other linguistic authenticity prevails in Rosny's Oulhamr, or in the successive groups of Xuah, Ourham, Wurehwê, and Wéréhé in Pelot's multi-epoch saga. These latter, moreover, appear to be a representation of phonetic change, though the shifts cannot be held to be clear cases of simplification or complexification.

---

5. A reviewer observes that the forename *Ernest* may in fact be semantically charged here, as it is in Oscar Wilde's play *The Importance of Being Earnest*.

## 7. Politeness as complexity

The final area we will consider is that of politeness and face-saving. Here, to put it simply, it appears that there are practically no politeness markers at all in prehistoric fiction. The standard request form is the bare imperative, and the words “please”, “thank you”, “sorry”, “excuse me”, “forgive me” or the request forms “can you” or “could you” are absent or vanishingly rare in the corpus. The absence of “thank you” even merits a metalinguistic clarification in Auel’s *The Clan of the Cave Bear*:

The Clan had no term for “thank you” [...]. They understood gratitude, but that carried a different connotation, generally a sense of obligation, usually from a person of lesser status. They helped each other because it was their way of life, their duty, necessary for survival, and no thanks were expected or received. Special gifts or favors carried the onus of obligation to return them with something of like value; this was understood and no thanks were necessary. (Auel 1980, chap. 13)

Thanking is shown to be absent *linguistically*, gratitude being expressed by actions emerging from a sense of obligation or by gift giving. Despite this frame, however, there is a moment later in the novel, where one character perceives a problematic lexical gap:

Ayla [...] wanted to express a feeling that was akin to gratitude, but not the way gratitude was normally felt, not gratitude that carried a sense of obligation or the kind a woman usually expressed to a man. She wanted to say something to him as a person, she wanted to tell him she understood. *She wanted to say thank you*, thank you for giving me a chance, but she didn’t quite know how.

“Brun, this girl is... is grateful to you. You said that to me. You said you were grateful for Brac’s life. *I am grateful to you* for my own.” (Auel 1980, chap. 17, my italics)

There are in fact expressions of thanks among Auel’s clan members. The “I am grateful to you” here is one example, but in another passage the praxis insists even more strongly on the performativity of the utterance:

He straightened and with the most formal of gestures, made so everyone could see, he added, “The leader of this clan is grateful to the girl for saving the life of the son of the mate of the son of my mate.” (*ibid.*, chap. 16)

The issue in fact appears not so much the distinction between gratitude in kind and gratitude in language, but a perceived complexity incorporating performativity *and* the formulaic nature of the sentence-word “thank you”. The terms “grateful” and “gratitude” occur frequently in the novel in narrative and in direct speech (as in the passages quoted above), to describe the behavior of others, while “thank you” is absent and the verb *thank* has only two occurrences outside the above quotation<sup>6</sup>.

We find a similar transfer of gratitude towards actions (announced linguistically) in *Rahan*. In one image Rahan announces his future intention to show thanks through the action of capturing a stag (Figure 3).



Figure 3: *Rahan*, n° 24, p. 36.

A second example is immediately performative, yet the verb *remercier* is preferred to the word-sentence *merci*, and further a third-person pronoun is used self-referentially by the speaker (Figure 4).

6. Each has two coreferent tokens: “She paused for a moment, clutched her amulet, and thanked the spirits for willow. She always thanked the spirits for willow, for its ubiquitous presence as well as for its painkilling bark”; “a spring feast [...] to give thanks to the spirits of the clan’s protective totems for seeing them through another winter. Ayla felt she had special reason to give thanks to her totem” (1980, chap. 3, 12).

Such avoidance of the formulaic “thank you” or “thanks” appears to be a defining language feature. While the preference for non-formulaic uses of the verb *thank* represents a complexity (in terms of number of words for instance), the avoidance of the sentence-word may also be seen as a symbolic promotion of less evolved forms of syntax.



Figure 4: *Rahan*, n° 24, p. 37.

## 8. Conclusion

The prehistoric fiction genre is marked by a range of representations of language simplicity, which creates a homology for the presumed comparative simplicity of the life of prehistoric people. This is manifested in several ways, in the works of novelists and comic-book authors. The replacement of personal deixis in pronouns by delocutive forms is notable (*Rahan, La guerre du feu*) though not systematically employed. Tense and voice may show restrictions associated with the here-and-now of the narrative (*The Inheritors*). The naming of referents such as animals, natural features, and tools in the prehistoric environment may be of a descriptive type (*Rahan, La guerre du feu*), and this may extend to names of humans and of their tribes (*Clan of the Cave Bear, Rahan*). Names of people may feature both phonetic simplicity, and orthographic complexity.

Finally, recourse to politeness strategies involving sentence-words or attenuating modalization is practically absent; those expressions of gratitude that there are use full syntactic structures rather than formulaic terms such as “please” and “thank you”.

## References

### Academic publications

- Ballard, Michel. 2001. *Le nom propre en traduction*. Gap & Paris: Ophrys.
- Cerquiglioni, Bernard. 1995. *L’accent du souvenir*. Paris: Minuit.
- DePaolo, Charles. 2000. Wells, Golding, and Auel: Representing the Neanderthal. *Science Fiction Studies* 27(3). 418-438.
- Everett, Daniel. 2009. *Don’t Speak, There Are Snakes: Life and Language in the Amazonian Jungle*. London: Profile Books.
- Li, Xueqin, Garman Harbottle, Juzhong Zhang, & Changsui Wang. 2003. The earliest writing? Sign use in the seventh millennium BC at Jiahu, Henan Province, China. *Antiquity* 77. 31-44.
- Mandala, Susan. 2018. “Listening” to the Neanderthals in William Golding’s *The Inheritors*: A sociopragmatic approach to fictional dialogue. *Discourse, Context & Media* 23. 62-69. DOI: <https://doi.org/10.1016/j.dcm.2017.05.006>
- Ruddick, Nicholas. 2009. *The Fire in the Stone: Prehistoric Fiction from Charles Darwin to Jean M. Auel*. Middletown, Connecticut: Wesleyan University Press.
- Tallerman, Maggie & Kathleen R. Gibson, eds. 2013. *The Oxford Handbook of Language Evolution*. Oxford: Oxford University Press.

### Websites (consulted 6 March 2023)

- Le lexique de Craô. *Rahan.org*. [www.rahan.org/presentation/lexicrao.html](http://www.rahan.org/presentation/lexicrao.html)
- Les Romains de la Préhistoire*. [www.lesromansdelaprehistoire.com](http://www.lesromansdelaprehistoire.com).
- Rahan.org. [www.rahan.org](http://www.rahan.org).

### Works of prehistoric fiction

- Auel, Jean M. 1980. *Clan of the Cave Bear*. London: Hodder & Stoughton.
- Berthet, Élie. 1876. *Romans préhistoriques*. Paris: E. Dentu.
- Cénac, Claude. 1967. *Les cavernes de la rivière rouge*. Paris: Magnard.

- Cranile [Arcelin], Adrien. 1872. *Solutré, ou les Chasseurs de Rennes de la France centrale: Histoire préhistorique*. Paris: Hachette.
- Golding, William. 1955. *The Inheritors*. London: Faber & Faber.
- Klapczynski, Marc. 2003-2010. *L'odyssée du dernier Neandertal*. Paris: Aubéron. [Series of three novels : *Aô l'homme ancien* (2003), *Le pouvoir d'Iktia* (2006), *Tsinaka, l'oeil de la toundra* (2010)].
- Lécureux, R. (text) & A. Chéret (artwork). 1969 (and after). *Rahan, fils des âges farouches*. [Many publications and collections.]
- Lewis, Roy. 1993. *The Evolution Man, or, How I Ate My Father*. New York: Vintage [orig. publ. as *What We Did to Father*, London: Hutchinson, 1960].
- London, Jack. 1906-1907. *Before Adam*, serialized in *Everybody's Magazine*.
- Pelot, Pierre. 1996-2001. *Sous le vent du monde*. Paris : Denoël. [Series of five novels : *Sous le vent du monde* (1996), *Le nom perdu du soleil* (1998), *Debout dans le ventre blanc du silence* (1999), *Avant la fin du ciel* (2000), *Ceux qui parlent au bord de la pierre* (2001)].
- Rosny, J.-H. ainé. 1911. *La guerre du feu*. Paris: Fasquelle.

### Other works of fiction cited

- McCarthy, Cormac. 1993. *All The Pretty Horses*. New York: Knopf.
- Theroux, Paul. 1980. *The Old Patagonian Express*. Harmondsworth: Penguin.



Comparer, niveler



# Complexity or *copia*? Latin versus Greek in the sixteenth century<sup>1</sup>

**Raf Van Rooy**

KU Leuven

---

## ABSTRACT

In this paper, I argue that sixteenth-century grammarians did not operate with a concept of linguistic complexity. Instead, in their intuitive comparisons of languages, they started from the widespread humanist concept of *copia*, “abundance”. Focusing on contrastive discussions of Latin and Greek, I aim to illustrate how the *copia* concept structured their assessment of these and other languages, especially the European vernaculars. In Latin-versus-Greek contrastive descriptions, found especially often in Greek grammars of the time, Greek usually emerged as the most copious language due to such “extra” categories as the article, dual, aorist, and middle voice. There were, however, respects in which Latin enjoyed more *copia*, such as certain nominal forms of the verb and the ablative case. Finally, I offer some reflec-

---

1. This article is an outcome of my research as a junior postdoctoral fellow of the Research Foundation – Flanders (FWO) [project 12V4818N] at KU Leuven, finetuned during my FWO senior postdoctoral fellowship [project 12V4821N] at the same institution, and during my MSCA Individual Fellowship, funded by the Horizon 2020 program under the Marie Skłodowska-Curie grant agreement No 890397, MigraLing, hosted at the Department of Philosophy, Classics, History of Art and Ideas (University of Oslo). I thank the two anonymous reviewers for their exceptionally kind and constructive criticisms, which have greatly improved the contents and structure of this paper. To Luz Van den Bruel I owe several stylistic improvements in my English. All remaining errors are mine alone.

tions on the relationship of the *copia* concept to modern notions of “complexity” and “difficulty”, arguing for the importance of factoring in cultural circumstances in complexity research.

#### KEYWORDS

complexity, abundance, *copia*, Latin, Greek, 16th century

#### RÉSUMÉ

Dans cette contribution, nous tentons de démontrer que les grammairiens du seizième siècle ne recourèrent pas à un concept de complexité linguistique. En revanche, leurs comparaisons intuitives des langues eurent comme point de départ le concept humaniste bien connu de *copia* “abondance”. En nous fondant sur des remarques contrastives relatives aux langues latine et grecque, nous essayons de montrer comment le concept de *copia* structura leur appréciation de ces langues ainsi que d’autres, en particulier les langues vernaculaires d’Europe. Dans ces descriptions contrastives opposant le latin au grec, qui apparaissent surtout dans les grammaires grecques de l’époque, le grec se manifeste généralement comme la langue la plus abondante, étant pourvue de catégories “extra” telles que l’article, le duel, l’aoriste et la voix moyenne. Le latin, quant à lui, possède quelques traits qui contribuent à sa *copia*, comme certaines formes nominales du verbe et l’ablatif. Enfin, nous proposons quelques réflexions sur les liens que *copia* entretient avec des notions modernes comme “complexité” et “difficulté”, tout en insistant sur l’importance d’intégrer les circonstances culturelles dans l’étude de la complexité linguistique.

#### MOTS-CLÉS

complexité, abondance, *copia*, latin, grec, XVI<sup>e</sup> siècle

## 1. Introduction: Complexity, a disparate modern concept

Historiographic studies of linguistic complexity tend to focus on the modern period, especially the nineteenth century, when most scholars agreed that languages exhibited different degrees of complexity, and the twentieth century, when many linguists came to agree on the opposite viewpoint, that “all languages are equally complex” (see Joseph & Newmeyer 2012 for a concise survey). Research conducted in recent decades has, however, overhauled the general validity of this maxim. The papers in Miestamo *et al.* (2008), for instance, showcase

very diverse perspectives on the matter, with some researchers still adhering to the equalist idea, arguing that there is a compensation mechanism in linguistic complexity, while others refute it. At the moment, the pendulum seems to be swinging more and more to the idea that there are differences in complexity that can be measured (cf. the contributions in Newmeyer & Preston 2014).

However, the results of current scholarship are tentative at best, since linguists have not yet come to a common definition of complexity. The general philosophical definition provided by Rescher (1998: 1) has nonetheless inspired numerous researchers since Miestamo *et al.* (2008), who can be credited with bringing it to linguists' attention:

Complexity is first and foremost a matter of the number and variety of an item's constituent elements and of the elaborateness of their interrelational structure, be it organizational or operational.

Applied to language, this definition leads to a very general concept of linguistic complexity, which I adopt here. Since there is no consensus definition of linguistic complexity in present-day scholarship, for my purposes it will suffice to work with this general concept, which has the advantage of covering most modern interpretations of linguistic complexity<sup>2</sup>.

The main hypothesis of my contribution will be that this modern concept, in all of its diversity, was mostly alien to earlier language studies, especially those of the sixteenth century, when linguistic diversity and difference first started to attract scholarly attention. I argue instead that sixteenth-century language enthusiasts operated with a different, though cognate concept, namely *copia*, "abundance", which grammarians borrowed from the sister discipline of rhetoric. Renaissance humanists valued *copia* highly, as my analysis will endeavor to show, starting from observations on properties of Ancient Greek in contrast to those of Latin. In particular, I will try to illustrate that humanists can be said to have followed the maxim

---

2. Cf. Rescher (1998: 9) on the different modes of complexity, applied to linguistics by Miestamo *et al.* (2008: VIII-IX). In Section 4, I will also briefly touch on another major approach to linguistic complexity, i.e. the idea that complexity should be measured in terms of difficulty and cost for language users and learners (see Miestamo *et al.* 2008: x).

“some languages are grammatically more copious than others”, thus adopting a non-egalitarian attitude to linguistic diversity.

My argumentation will start with a contextual outline of Latin-versus-Greek discussions against the backdrop of the ever-widening language horizon of sixteenth-century Europeans (Section 2). Then I will discuss the alleged abundance of Greek in comparison to Latin as the idea emerges in a variety of grammatical sources from the sixteenth century, mainly Greek grammars but also handbooks on Latin and the vernaculars (Section 3). My discussion of this theme will be arranged according to the grammatical categories in which Greek is claimed to exhibit abundance, or a lack thereof (Sections 3.1-3.6). Finally, I will discuss the position of the *copia* concept in the Renaissance grammaticographic landscape and its relationship to the concepts of complexity and difficulty (Section 4).

## 2. Latin versus Greek in multilingual Europe

*Articulus nobis nullus, & Græcis superfluous, nisi cum rem notam repetit subijcere intellectui. at tunc est relatiuum. Idem enim est, ὁ δοῦλος μου, quod δοῦλος ὃν οἶσθα: alioqui otiosum loquacissimæ gentis instrumentum est*<sup>3</sup>.

‘We have no article, and for the Greeks it is superfluous, except when it repeats a known matter in order to make it subject to understanding. But in that case it is a relative. For *ho doûlós mou* (“the slave of me; my slave”) is the same as *doûlos hòn oîst<sup>h</sup>a* (“slave whom you know”); otherwise it is an idle instrument of an excessively chatty people.’

The Italian humanist Julius Caesar Scaliger (1484–1558) famously discarded the Greek definite article as the “idle instrument of an excessively chatty people” in his widely read *De causis linguæ Latinae*<sup>4</sup>. The feature made the language needlessly elaborate, and the category of the article itself was superfluous (Latin: *superfluous*), since it was not needed to understand the Latin language, which Scaliger,

---

3. Scaliger (1540: 126). I quote sixteenth-century works in their original orthography, resolving abbreviations between square brackets.

4. For a recent modern edition, see Scaliger (2019). On the article quip and its quirky fate, see the appendix in Van Rooy (2021). All English translations in this contribution are mine.

though a talented Hellenist, regarded as the most important tongue of all. The article was, in other words, a linguistic overcomplication. Scaliger's verdict on the Greek dual number was no less adverse. This grammatical frivolity, used to refer to exactly two entities as opposed to the singular and plural numbers, was typical of the Ionic dialect of Greek, claimed Scaliger (1540: 139), who called it "*nugacitas illa Ionum*" ("that futility of the Ionians"). This Ionic dual was moreover deficient, lacking as it did full-fledged conjugations and declensions for all persons and cases. Fortunately, the Aeolic dialect did not retain this feature; hence the dual was not adopted by the Romans<sup>5</sup>.

Scaliger, although adhering to the traditional idea of Aeolism, which implied that Latin derived from Aeolic Greek, made no secret of his predilection for the language and literature of ancient Rome at the expense of those of classical Greece<sup>6</sup>. Not only did he prefer Vergil to Homer and Seneca to the Greek dramatists, but he also sought out features in which the Latin tongue excelled Greek,

---

5. Scaliger (1540: 139): "*Quare Iones non recte fecere, qui Dualem numerum à plurali discernere: atq[ue] iccirco seueriores Aeoles neq[ue] recepere, neq[ue] in Latinos transmisere. & nugacitas illa Ionum in multis temporibus uerboru[m] personas aliquot non potuit eruere in eo numero: in nominibus autem pauculos casus expressere.*" Scaliger's mistaken idea that the dual was an Ionic rather than an Attic or Homeric feature, and the view that Aeolic and hence Latin had lost this number, was taken from the Greek dialectological canon of the time, consisting of a collection of texts attributed to Plutarch, John the Grammarian (John Philoponus), and Gregory of Corinth (see the appendix in Trovato 1984). On the dual as an Ionic feature, see John the Grammarian in Lascaris & Manutius (1512, sig. ξ vi<sup>v</sup> / x vii<sup>v</sup>): Η δὲ τῶν δυϊκῶν ἀκρίβεια παρ' αὐτοῖς, φυλάσσεται. / "*Dualium autem diligentia apud eos seruat*ur." On the absence of the dual in Aeolic, see Gregory of Corinth in Lascaris & Manutius (1512, sig. ρ ii<sup>r</sup> / & i<sup>v</sup>): Τοῖς δυϊκοῖς ἀριθμοῖς οὗτοι κέχρηται οὐδαμῶς, καθὰ δὴ κ[αί] οἱ ῥωμαῖοι τούτων ὄντες ἄποικοι. / "*Dualibus numeris hi utunt[ur] minime, quemadmodum & romani, qui ex his sunt deducti.*" I quote from a sixteenth-century edition with contemporary Latin translation since no up-to-date modern critical editions of this text corpus are available or readily consultable. Scaliger's qualification of the Aeolians as "seuiores" ("more austere") might echo Martial's complaint that the Latin muses were sterner than the Greeks' in *Epigrammata*, 9.11.16-17: "*Nobis non licet esse tam disertis, / Qui Musas colimus seuiores.*"

6. For ancient interpretations of Aeolism, see Stevens (2006-2007). For early modern conceptions and transformations of Aeolism, see Van Rooy (2020a: 78-80).

while at the same time granting that Latin grammar was indebted to Greek in some respects, such as the accent and the comparative (Scaliger 1540: 95 & 206). Scaliger even went so far as to advance that Latin had certain periphrastic constructions reflecting the Greek aorist, traditionally regarded as a tense which Latin lacked. For instance, the verb form *caesus fuit*, “he was beaten”, was the Latin aorist pendant of Greek ἐτύφθη (*etúp<sup>h</sup>t<sup>h</sup>ē*), whereas *caesus est*, “he is beaten”, corresponded to the Greek perfect tense form τέτυπται (*tétuptai*). Scaliger probably had no aspectual nuances in mind when discussing this subtle difference, but rather understood it according to the traditional Stoic idea that the aorist referred to an event occurring at an “indefinite” moment in the past, whereas the perfect concerned an event in the recent past (Lallot 2013).

Scaliger, in sum, consciously pitted Latin against Greek in his rationalized account of Latin grammar, a linguistic battle in which the former gained the upper hand. It is striking that he regarded features such as the Greek article and dual as useless phenomena needlessly complicating the grammar of this language. Scaliger’s case seems to suggest that excessive complexity was not considered a linguistic asset in the sixteenth century; instead, one should strive for efficiency and rationality in grammar. Was the Italian humanist voicing a common opinion, or was he reacting against established ideas? I will argue that he was indeed denouncing contemporary views and that his reaction should be framed in a debate not on linguistic complexity but rather on linguistic abundance. Before embarking on a detailed analysis of how the discussion on *copia* in Greek versus Latin took place, I will elaborate on the setting in which this contrastive approach to grammar thrived.

Whereas Latin had been the main subject of linguistic study in western Europe in the Middle Ages, scholars had started to study Ancient Greek in depth only very recently, guided by learned Byzantines who fled their crumbling empire<sup>7</sup>. The main motivation was to better understand ancient Roman literature, whose study required a shift of attention from scholastic to classical Latin,

---

7. On the so-called rediscovery of Ancient Greek in Renaissance Italy and beyond, see e.g. Saladin (2000), to be read alongside the critical review by Pontani (2002).

and later also the New Testament, originally composed in Koine Greek. However, Greek and Latin were not the only languages that occupied scholars' minds at that time. The scope of Renaissance language studies also covered the vernacular tongues. After initial developments during the later Middle Ages, the standardization of these languages gained momentum in this period, partly due to the commercialization of the printing press. These vernaculars were increasingly codified and uniformized in grammars based on classical models, usually Latin but sometimes also Greek. In addition to the classical and vernacular tongues, Hebrew, too, attracted great interest as the language of the Old Testament and was often regarded as the primeval tongue. Scholars also developed an interest in extinct ancient languages such as Gaulish – so-called *Trümmersprachen* – as well as in the languages of other continents, which they mainly described in the form of missionary grammars<sup>8</sup>.

This boom in grammaticography started from more or less the same blueprint for all languages, namely the traditional parts-of-speech model developed by the ancient Greeks and adapted by Roman grammarians. This state of affairs generated a kind of universal touchstone to judge a language by. But no clear concept of language complexity as defined in broad terms in Section 1 played a central role in sixteenth-century assessments of individual languages. Indeed, language complexity in this sense seems to be a very modern concept, marginal at best in the sixteenth century. I contend instead that a general notion of *copia*, “abundance”, typically informed language assessments in the Renaissance, and that contrasting observations on Latin and Greek can convincingly make this point. I will draw on a variety of Latin and Greek grammars from the sixteenth century in deploying my argument, paying attention to both the contents of the contrastive observations and the terminology typically used to make such remarks.

How was Greek contrasted with Latin by sixteenth-century scholars? Before trying to answer this central question, I first have to highlight in which contexts and for what purposes such contrastive statements were made. These observations tend to

---

8. For a more detailed sketch of the Renaissance linguistic landscape, see Van Rooy (2020b: 49-52) and the references there.

occur mainly in grammars of Ancient Greek due to two main interrelated circumstances. With very few exceptions, students started out with Latin before embarking on their study of Greek. Latin was also the default metalanguage used to describe Greek in the sixteenth century, after an initial trend of monolingual Greek grammars (see Botley 2010). These two facts made it natural for teachers and grammarians to rely on their audience's knowledge of Latin to show where the two classical languages overlapped and where they differed. The other way around made less sense, since offering extra information on Greek grammar to students of Latin would be an unnecessary overcomplication in their early learning career. Moreover, Latin grammars were either monolingual Latin or bilingual Latin-vernacular, a situation obviously not favorable for making contrastive observations on Latin versus Greek<sup>9</sup>.

The sixteenth-century tradition of putting Greek next to Latin again can be viewed as a kind of reversal of what had happened in antiquity, when Latin grammar evolved out of its Greek model, and certain Latin features came to be put next to (near-)equivalent phenomena in Greek (see Van Rooy forthcoming). This ancient comparison led to the *locus communis* that Latin was poor compared to Greek, especially in terms of scientific and philosophical jargon, and at best a barbaric descendant of (Aeolic) Greek, an inferiority complex that would haunt the language until the early modern period (Fögen 2000). Some Roman writers did, however, try to formulate this *topos* in a more positive fashion by accusing the Greeks of verbosity (Fögen 2000: 49), an accusation Scaliger also leveled. Authors from (late) antiquity tended to focus on the lexicon when assessing the relative richness of the two classical languages. A major exception was the rhetorician Quintilian, who also paid attention to grammatical features of Latin versus Greek, a lead followed by sixteenth-century humanists, as I set out to demonstrate now.

---

9. The Latin-vernacular context would also have been a good starting point for conducting a similar analysis, but my research interests drew me to observations on Latin vs. Greek. Perhaps a follow-up study could focus on the Latin-vernacular context to see if the results would be similar, and I offer some pointers throughout Section 3 below, especially in Sections 3.1 & 3.6.

### 3. The *copia* of Greek

Which features are prototypically discussed when highlighting the differences between the two classical languages? Scaliger's remarks, rather atypically made in a discussion of Latin rather than Greek grammar, already show in which direction to look for the most prominent features contrasted: the article, the dual, and the aorist, in addition to various nominal forms of the verb, the middle voice, and the ablative.

#### 3. 1. The article

Latin has no definite article, as Greek and many vernaculars in western Europe have, a fact eagerly underlined by grammarians of both Greek and vernacular languages. Typically, they used the Latin verb *carere*, "to lack", to point out that this language is deficient in this regard. Philipp Melanchthon, the Protestant Hellenist who wrote popular handbooks of both classical languages, for instance observed in his 1526 Latin grammar that "*latinus sermo in uniuersum caret articulis, Germani habent articulos, quemadmodum & græci*" ("the Latin language entirely lacks articles; the Germans have articles, just as the Greeks do, too"; Melanchthon 1526, sig. F 7<sup>r</sup>).

In an outline of Greek syntax published a year later, the Latin lack of an article is likewise stressed in similar terms:

*Latini omnino carent. In peregrinis idiomatibus, vti Gallico seu germanico inuenire licet quippiam, articuli vice positum. Nam cum ὀβασιλεὺς [sic] ἦλθε, Profero, latine sonat Rex venit, Articuli virtus non explicatur.* (Winter 1527, sig. c ij<sup>r-v</sup>)

'The Latins entirely lack [the article]. In foreign idioms, such as French or German, one can find something put in the role of the article. For when I say *ho basileûs êlt<sup>h</sup>e* ["the king came"], in Latin it is *rex venit*; the value of the article is not made explicit.'

Apart from using the same verb *carere*, the grammarian Johann Wint(h)er von Andernach mentioned that Latin could not express the *virtus* of the article, a term which I have rendered neutrally as "value" but which often has a distinctly positive connotation in Latin. At the risk of overinterpretation, one might even suggest that Winter saw the possession of an article as an indisputable linguistic quality,

a token of *copia*. Additional observations similar to Melanchthon's and Winter's could easily be noted, including examples from grammars of other languages, such as French or Hebrew<sup>10</sup>. The recurring terminology seems to indicate that Latin was regarded as deficient in this respect and was surpassed by Greek, the vernaculars, and Hebrew. This harsh judgment of Latin is all the more remarkable since the Latin grammatical tradition had nonetheless borrowed the article concept from Greek tradition but had applied it to demonstrative pronouns serving as case and gender markers in tables outlining Latin nominal morphology (see Denecker & Swiggers 2018). Humanists, however, unmasked this so-called Latin article, which was especially prominent in medieval grammars of the language, as a pronoun.

### 3. 2. The dual

Latin also lacked the dual number, a fact which Scaliger applauded but many Greek grammarians saw as a deficiency. Melanchthon notably described the Greek number system as follows in the revised edition of his Greek grammar manual of 1520:

*Nu[m]eri tres, Singularis. Dualis & pluralis. Vincit e[ni]m latinu[m] sermone[m] duali nu[m]ero græca copia.* (Melanchthon 1520, sig. B iii<sup>v</sup>)

'There are three numbers: singular, dual, and plural, since the Greek abundance defeats the Latin language by the dual number.'

Remarkable, here, is the terminology used by Melanchthon to chart the relationship between Greek and Latin. Not only did the Hellenist narrowly associate Greek with the crucial term of *copia*, suggesting that Latin was not as abundant, but he frankly put forward that Greek bested (*vincere*) Latin on account of the dual number. Again, no mention is made of a notion of complexity or the like but, very explicitly, of abundance.

---

10. For a French grammar, see Cauchie (1576: 72): "*Articulum quo Latini carent, cum Grecis Germanisq[ue] communem habemus.*" For a Hebrew grammar, see [Restauld de Caligny] (1533: 29-30): "*Quam habeant significationem Articuli, Latine non potest proprie exprimi: nam Latini carent Articulis. Explicabimus igitur Græcè & Gallicè.*" On the article category from ancient Greece to the Greek revival during the Renaissance, see Matthaios & Van Rooy (2021).

### 3. 3. The aorist

In contrast to Scaliger, most sixteenth-century scholars agreed that Latin also lacked the aorist, a Greek verb form expressing aspect rather than tense, a fact poorly understood before modern linguistics. Indeed, the aorist was traditionally categorized as a grammatical tense, for instance by the humanist grammarian of Greek Ludwig Lopad, an obscure Protestant scholar whose Greek manual has barely been studied. Lopad's remark on the aorist can be neatly read in conjunction with that of his reformed colleague Melanchthon on the dual number:

*Nomenclatura temporu[m] Latini uincimur à Græcis, significatio[n]e no[n] ite[m]. Cæteru[m] paraciminu[m] ab Aristo. [sic] sic discernemus, na[m] hæc duo tempora sæpiuscule auditoru[m] animos obturba[n]t, παρακειμενος aut[em] significatio[n]e[m] semper præteriti perfecti seruat. ἄριστος in orationis co[n]textu, interdu[m] imperfecti, interdu[m] perfecti, interdum plusquamperfecti retinet significatum. nemo hæc facile aßequetur, nisi stili exercitio: hic enim rectißimus artium perfector est, teste Cicerone. (Lopad 1536, f. 23')*

'We Latins are defeated by the Greeks in tense terminology but not likewise in their meaning. Still, we will distinguish the *parakeímenos* from the aorist as follows, for these two tenses rather often confuse students' minds. The *parakeímenos* always preserves the meaning of the past perfect; the *aóristos* retains in the speech context sometimes the meaning of the imperfect, sometimes of the perfect, sometimes of the pluperfect. No one will easily master these, except by stylistic exercise, for this is the best method to perfect the arts, as Cicero testifies.'

Like Melanchthon, Lopad suggested that the Greeks had the upper hand, this time in the realm of verb tenses. Yet the grammarian was not unyieldingly positive about this Greek extra, for the meaning of the aorist was rather fuzzy, or undetermined as its name indicates, belonging to the domain of stylistics rather than grammar, and thus in a position to easily confuse students. The Greek aorist was therefore not an indisputable asset, and Lopad thus advanced a more nuanced idea about the *copia* of Greek vis-à-vis Latin. Perhaps this reservation has to do with the fact that Lopad sided with Latin speakers (cf. "*Latini uincimur*"). Be that as it may, the Hellenist strikingly resorted to the same Latin verb as Melanchthon, *vincere*,

“to defeat”, in sketching the balance of power between both classical languages, here struck in favor of Greek but not unequivocally so.

### 3. 4. Nominal verb forms: Participles vs. gerund and supine

Another grammatical feature which Latin lacked and Greek possessed was a wealth of participial forms, including an active perfect participle and, obviously, the gamut of aorist participles. These Greek forms needed to be paraphrased in Latin, as the Swiss Hellenist Jacobus Ceperinus (ca. 1500–1525) observantly noted in his 1522 Greek grammar:

*Latina lingua participiis præteritis actiuis carente, hæc græca participia præterita per relatiuu[m] & uerbum præteritum circumscribimus, ut ὁ τετυφῶς τύψας τυπῶν qui uerberauit, τοῦ τετυφότηος eius qui uerberauit &c. (Ceperinus 1522, sig. C vii<sup>v</sup>)*

‘Since the Latin language lacks active past participles, we paraphrase these Greek past participles by means of a relative [pronoun] and a past verb, as *ho tetup<sup>h</sup>ōs túpsas tupōn*, “he who beat”, *toū tetup<sup>h</sup>ōtos* “of him who beat” etc.’

The deficiency of Latin is once again stressed here by means of the verb *carere*, “to lack”, while the grammarian here, too, identifies himself with speakers of Latin rather than with those of the more copious Greek language. It seems, therefore, that several Hellenists educated in the Latin-centered west refrained from promoting their own scientific lingua franca, Latin, even admitting instead that Latin was inferior to Greek. In other words, their admiration for ancient Greek culture and its language was stronger than their love for the classical Latin tongue, their preferred means of communication.

Grammarians did recognize that Latin possessed a number of nominal verb forms which Greek lacked, the gerund and the supine. In the widely used Jesuit grammar of Greek by Jakob Gretser (1562–1625), first published in 1593, the Greek alternatives for these nominal verb forms are discussed as follows:

*GERVNDIIS & supinis carent Græci. Eorum loco vtuntur uerbis infinitiui modi, interdum cum præpositione & articulo, interdum sine his [...]. (Gretser 1593: 155)*

‘The Greeks lack gerunds and supines. In their stead they use verbs in the infinitive mood, sometimes with a preposition and article, sometimes without them.’

Notably, the terminology to mark deficiencies of Latin, i.e. the verb *carere*, is here applied in a similar vein to Greek, suggesting that the Greek *copia*, too, had its limits, and that in some respects Latin was considered more abundant.

### 3. 5. The middle voice

Several humanist grammarians took one other attribute of the Greek verbal system as a source of linguistic abundance in contrast to Latin. The Protestant Hellenist mentioned above, Ludwig Lopad, provides a revealing remark on this subject:

*Qui aute[m], inquis uoces mediæ à passiuis dignoscuntur? respondeo, uerborum structura mo[n]strabit: possunt aut[em] ab omnibus uerbis trahi media, in causa est huius linguæ ubertas in præsentia uolo pueros redire in memoriam eorum, quæ supra diximus de medijs nempe quæ ab Actiuis ueniant, actiue et passiue significant. In summa omne uerbum medium retinet significationem sui primitiui: nullum uero uerbum Latinis est quod ego sciam, huic quod respondeat. (Lopad 1536, f. 24<sup>r-v</sup> [bis])*

‘But how, you say, are middle voices to be distinguished from passive? My reply: the structure of words will show it. The middle can be drawn from all verbs, because of the abundance of this language. Presently, I want boys to remember what we have said above on middle verbs, namely those which come from active verbs, and have an active and passive meaning. In sum, every middle verb retains the meaning of its original root, but there is no verb among the Latins which, as far as I know, corresponds to this.’

Lopad stressed once again the abundance of Greek, here termed with *ubertas*, a synonym of *copia*. Unlike Latin, the Greek language has a verb form that can express activeness and passiveness at the same time, as, for instance, in the form *λούομαι* (*louómai*), “I wash myself”, where the subject of the verb is at the same time its object.

### 3. 6. The ablative

Lopad was aware that Latin, too, had something of its own, which Greek lacked. In fact, when outlining the Greek middle voice, he linked it to the Latin ablative. Both features were language-specific categories which, despite their particularity, were eagerly used in both writing and speaking the respective languages<sup>11</sup>. The ablative is often presented as one of the few aspects in which Latin is richer than Greek, even by grammarians who generally painted a more positive picture of Greek or who described languages other than Latin or Greek. Melanchthon, for instance, dryly noted in his 1526 Latin grammar that:

*Estq[ue] hic casus [sc. ablativus] latinis proprius græci.n. ablatiuos non habent. Ideo grammatici no[n]nunqua[m] latinum casum uocant.* (Melanchthon 1526, sig. C 3<sup>r</sup>)

‘This [ablative] case is proper to the Latins, for the Greeks do not have ablatives. For this reason, the grammarians sometimes call it the Latin case.’

Melanchthon’s terminology for this Greek shortcoming was, however, less powerful than his repeated descriptions of what Latin “lacked” (*carere*) or in which respects Latin was “defeated” (*vincere*) by Greek. Melanchthon used instead the more neutral “not to have” (*non habere*). The Jesuit grammarian Gretser (1593: 14) resorted to stronger language when stating, very succinctly, that “the Greeks lack the ablative” (“Ablatiuo carent Græci”), thus using terminology similar to that used to describe Latin defects.

Some grammarians seem to have believed that having no ablative was not a deficiency but rather a positive feature. An interesting example can be read in the French grammar of Antoine Cauchie (1535–1600), who preferred to follow the Greek model as was fashionable in the sixteenth century, when describing the French noun declensions in terms of case: “*Ablativum Græcos secuti sine inco[m]modo rejicimus*” (“Following the Greeks, we reject the ablative

---

11. Lopad (1536, f. 24<sup>r</sup> [bis]): “*Nolo præterea pueros ignorare, media uerba esse familiarissima Græcis. quia sunt propria. Que[m] admodum Latinis familiaris est ablatiuus non alia causa nisi quia proprius: horum igitur magnus est usus scribendo & loquendo, nam propria [...]*”

without any inconvenience"; 1576: 84). Cauchie's remark suggests that the ablative was not a sign of Latin abundance but rather of redundancy. Lacking this case, Cauchie seems to have implied, did not hurt the *copia* of Greek nor, by association, French. Other humanists, most notably Philipp Nicodemus Frischlin, proposed a different solution to make up for this defect of Greek by stipulating an ablative case for this language as well<sup>12</sup>.

The example from Cauchie allows me to briefly make an additional point here. In sixteenth-century grammar, contrastive observations are surely not limited to Latin and Greek, and the discourse in terms of "lacking", "defeating", and "abundance" is not at all restricted to these two languages. In fact, numerous examples could be cited showing that vernacular grammarians were often seduced into describing the *copia* of their native tongues in distinctly positive terms in contrast to other vernaculars and the classical languages. Various aspects of the French verbal system were, for instance, applauded by Jacobus Sylvius (Dubois; 1478–1555) and Cauchie, who emphasized the poverty of either Latin or both classical languages in their appraisals of French abundance<sup>13</sup>.

#### 4. *Copia*, complexity, and difficulty

I can now return to my main point: did humanists have a concept of linguistic complexity, or did they work with *copia*, "abundance", which may be taken to partly overlap with complexity? I am inclined to adopt the latter view, taking the modern concept of linguistic complexity, for which the early moderns had no clear-

---

12. On Frischlin's daring view and his controversy with Martin Crusius on this subject, see Van Rooy (2021).

13. See Sylvius (1531: 118): "*Tempus nobis quotuplex Latinis est, sed non per omnes similiter modos. Nam imperatiui futurum, quod sciam, non exprimimus, nisi vt quibusdam videtur per voces praesentis, adiecta oratione hac, tempore futuro, le te[m]ps aduenir, ab aduenire, vel simili. Sed vt nos Latini eo vincant, vincuntur quoque à nobis vno praeterito.*" Note the verb *vincere* here. See Cauchie (1576: 149 [sic pro 139]): "*Qui ex hoc loco Latinis nos Graecisque jejuniore judicaverit, graviter erraverit, nam hic sumus sanè (quantum illorum mihi opes notæ sunt) multò opulentiores, cùm eorum passiva duplici praeterea ratione efferamus.*" Note here the contrasting comparatives *ieiunior* ("poorer") vs. *opulentior* ("richer").

cut terminological apparatus, as partly overlapping with *copia*, a typical Renaissance concept with roots in classical antiquity. Indeed, in order to avoid the risk of anachronism, it seems well-advised to frame sixteenth-century ideas by means of terms with which humanist scholars themselves operated, especially concerning such a prominent concept as *copia*. I will, however, also have to say a word on how I believe the modern concept of linguistic complexity relates to grammatical *copia*.

*Copia* was, beyond doubt, a positively loaded term in the eyes of Renaissance scholars. Yet *copia* was not originally a linguistic concept but a rhetorical one, celebrated by no one better than by Erasmus in his textbook *De copia*, first published in 1512 and with many later editions, the last authorized one appearing in 1534 (on this work, see e.g. Callahan 1978; Mack 2011: 80-88). *Copia* equaled richness but also rhetorical diversity, as Erasmus' (1534: 7) definition suggests: *copia* is *orationis variandae ratio*, "a method of varying speech" (cf. Diu & Vanautgaerden 2018). Hence variation, too, was a quality in humanist stylistics and, as my analysis above suggests, by extension also in humanist grammar. Erasmus promoted this conceptual transfer from stylistics to grammar himself, as he made suggestions for grammatical variation in Latin in *De copia* (especially book 1, chapters 34-94). Erasmus was led by the classical ideal of *copia* in many of his scholarly endeavors, his compilation of ancient *Adagia* being the most obvious substantiation of this interest. In formulating his *copia* concept, he was crucially indebted to classical Latin rhetoricians such as Cicero and especially Quintilian, who had entitled the beginning of book ten of his *Institutio oratoria* "De copia verborum" ("On the abundance of words"). Yet the Greek language and literature also occupied a central place in Erasmus' reflection on *copia*. Being very closely acquainted with Greek grammar, as he translated parts of Theodore Gaza's manual into Latin, the humanist from Rotterdam was convinced more than anyone of the abundance of this language, an opinion he voiced repeatedly in *De copia*<sup>14</sup>.

---

14. See Erasmus (1534: 12) for an early instance in this work: "Ad hæc, uertendis Græcis autoribus non mediocriter augebimus sermonis copiam, propterea quod hæc lingua rerum uerborumq[ue] ditissima est."

It was easy enough for humanists to extend their stylistic ideal of *copia*, lavishly depicted in many Renaissance art pieces as the cornucopia or horn of plenty, to grammar itself, given that grammar was considered the foundation of rhetoric. As such, rhetoric was in an excellent position to cross-fertilize ideas with grammar, especially since the strong logical orientation of the Middle Ages had disappeared in humanist milieus (Percival 1983). This left room again for the classical symbiosis of rhetoric and grammar, crystallized in Quintilian's *Institutio oratoria* (Ax 2011) and culminating in a new stylistic ideal of *copia* that reminisced the classical ideal of eloquence, or *copia dicendi* according to the *Rhetorica ad Herennium* (Mack 2011: 80). Not surprisingly, the cornucopia was variously employed as a metaphorical designation for collections of lexical and grammatical works, including, most famously, Niccolò Perotti's posthumously published commentary on Martial (Perotti 1489), in addition to a vast collection of Greek grammatical treatises issued by Aldo Manuzio's office in 1496 (Antinori *et al.* 1496). The idea of the grammatical cornucopia went beyond a mere fashion in titling, as I have tried to illustrate in Section 3. In sixteenth-century multilingual Europe, *copia* was transformed into a grammatical concept, grammar being the foundation of all studies in the education system of the time, and having become an increasingly diverse discipline covering not one but entire scores of different languages (see Section 2).

Might it be fruitful to take the image of the horn of plenty to try and visualize the ways in which sixteenth-century humanists conceived of *copia* as a grammatical concept? Languages were assessed against the background of the same grammaticographic model rooted in classical antiquity, initially based on Greek scholarship but soon tailored to Latin too (see Van Rooy forthcoming and the references there). This so-called Greco-Latin model can be regarded as the horn in itself, without its contents. Depending on the language described, this classical horn could be more or less filled, *copia* clearly being a gradational, or at least a relational, quality. The better the horn was filled with features matching or aligning with categories of the classical model, the more *copia* a language had. The horn's contents could, in other words, vary, whereas the model of the horn remained more or less stable. This visualization may, I believe, help explain

the terminology used in describing a language and its categories, especially the omnipresent verb *carere*, “to lack”, which humanists adopted from earlier Latin grammarians and tailored to their own linguistic descriptions, while attaching new connotations to these terms<sup>15</sup>.

Despite the obvious differences between the sixteenth-century rhetorical concept of *copia* and the modern linguistic concept of complexity, I believe it possible to find some overlap between them. *Copia* as applied to grammar referred to a high quantity of available forms and linguistic categories, as the numerous examples in Section 3 showcase; this concept of grammatical abundance coincides in part with what Miestamo *et al.* (2008: VIII) refer to as “compositional complexity”, one of two ontological modes of complexity after Rescher (1998: 9), covering:

- a. *Constitutional complexity*: number of constituent elements (such as phonemes, inflectional morphemes, derivational morphemes, lexemes).
- b. *Taxonomic complexity* (or *heterogeneity*): variety of constituent elements, i.e., number of different *kinds* of components (variety of phoneme types, secondary articulations, parts of speech, tense-mood-aspect categories, phrase types etc.).

I have focused on type b in this paper, whereas type a can be associated with the lexical abundance the humanists rightly insisted Greek had. As I have pointed out in Section 2, this latter idea, which I have not discussed in detail here, traces back to antiquity and was stressed in Erasmus’ *De copia*<sup>16</sup>. Despite this overlap, it seems ill-advised to use *copia* and complexity interchangeably when writing the history of complexity as a linguistic concept, if only because the contexts in which these concepts were elaborated differed greatly. *Copia* was the manifestation of a highly rhetorical view on grammar

---

15. The verb *carere* is, for instance, used by Priscian to designate the Latin lack of articles in his *Institutiones grammaticae* (Keil 1855-1880, vol. 3: 139): “*neque apud Graecos tamen pronomina demonstrativa, quippe praesentes ostendentia personas, quae in prima sunt cognitione, articulorum esse possunt capacia praepositivorum, quibus Latinitas caret, sicut docuimus.*”

16. See Erasmus (1534: 12), quoted in footnote 13.

thriving in a period of great linguistic competition, whereas the modern complexity debate in linguistics has been fueled by different factors, including cultural, political, and even racial arguments (Joseph & Newmeyer 2012: 361).

Needless to say, the *copia* concept was only one among several factors shaping sixteenth-century ideas on the hierarchy of languages. The sociolinguistic context of the time, for instance, also affected the *copia* ideal. Luce Giard (1992) has made a strong case for picturing the sixteenth-century linguistic scene as a battle among the vernaculars in which the classical languages and especially Greek were also involved, as my analysis seems to confirm<sup>17</sup>. The fact that humanist grammarians eagerly employed the Latin verb *vincere*, “to defeat”, in describing the *copia* of a language in contrast with another may be taken to indicate that abundance was a metaphorical weapon in the language battles of the sixteenth century. It would be worthwhile to take stock of all such weapons, which included apologies on behalf of individual languages and their codification in grammars, but this proposed endeavor exceeds the limits of the present contribution. I do want to venture some thoughts on a concept akin to complexity typically encountered in sixteenth-century apologies of languages, first formulated in favor of Greek and Hebrew but soon also of the vernaculars, the property of difficulty or facility of learning: how much effort does it take to learn a language? Greek provides an interesting case for this idea, too, as I will briefly illustrate by exploring a number of select apologies for this language.

Perhaps rather bafflingly from a present-day point of view, Greek was often promoted as an easy language, an observation that became a true *topos* in apologetic orations. The German humanist Petrus Mosellanus (1493–1524), a close friend of Melanchthon’s, emphasized the facility of Greek early (Mosellanus 1518, sig. E ij<sup>r</sup>). Yet this ease was usually not linked to intrinsic linguistic reasons but rather to previous knowledge of another language, i.e. Latin. If one already knew Latin, it was easy enough to learn Greek. This argument was

---

17. For the battle as it involved Greek in the Renaissance west, cf. also Saladin (2000) and the review by Pontani (2002), mentioned earlier in Section 2. See also Burke (2004: 17–25) for a substantial discussion of the widespread concern with the riches or poverty of a particular language in early modern times.

made by, among others, the Frisian Hellenist Suffridus Petri (1527–1597) in one of his five orations in praise of Greek, pronounced in Leuven but published in Basel<sup>18</sup>. Occasionally, humanists did promote certain features of Greek as simple, as, for instance, the German scholar Laurentius Sifanus (1510–1579) did in his 1558 oration on the language, in which he emphasized the morphological simplicity of Greek and the clarity of the article as gender marker<sup>19</sup>. The latter, in particular, seems to have been a fairly common idea. The Swiss Hellenist Martin Ruland the Elder (1532–1602) likewise regarded the article as a kind of simplifier of speech, connecting different parts of speech and making interrelationships transparent<sup>20</sup>. Here, we appear to be approaching a notion of complexity versus simplicity as became more common in later language studies. In this case, however, the complexity is not defined in the way I have done in Section 1 but by pedagogic criteria, in the sense that the complexity is determined by the ease or difficulty with which learners come to master a language (see Miestamo *et al.* 2008: x, xii & *passim*).

There is, in other words, a clear dichotomy in modern thought on linguistic complexity. According to one approach, complexity should be determined by an objective look at the language system; according to the other, complexity is situated in the mind of the language user and learner. The sixteenth-century situation already

---

18. Petri (1566: 67): “*Tanta enim eius descendæ facilitas est, ut nullam minori labore disci affirmare ausim, quàm Græcam: præsertim ab his, qui Latinam utcumque didicerunt: utpote quæ per omnia ferè cum Latina præceptis ipsis conuenit, & ipsas Latinorum regulas sui cognitione illustrat.*”

19. Sifanus (1564 [1558]: 20–21): “*Neq[ue] verò, tametsi o[mn]is ad virtutem, vt Hesiodus ait, ardua via est, ita difficile linguæ Græcæ studium est, vt vulgus putat. Primum enim articuli nullo negotio nobis genera nominum, vt in Germanico, ita in Græco sermone ostendunt. deinde quum casus & declinationes nominum, tum verborum tempora & coniugationes paucis ijsq[ue] certis regulis continentur. In Latinis verò nominibus & verbis varia[n]dis tanta cùm difficultas, tum incertitudo deprehe[n]ditur, vt plus negotij, plus laboris, plus fastidij in vna tertia Latina vel declinatione, vel coniugatione (vt de cæteris incommodis nihil dicam) percipienda sit, quàm in vniuersis Græcorum declinationibus coniugationibusq[ue] descendis inueniri potest.*”

20. Ruland (1556: 31–32): “*Illae [sc. partes orationes] enim per hos [articulos] componuntur, efficiunturque aptae ac intellectu faciles.*”

shows how unwieldy such a double concept of complexity is, since a humanist grammarian could at the same time judge Greek to be more copious and, to use modern terms, more complex than Latin in terms of language categories, yet still claim that it was easier to learn Greek than Latin. This apparent contradiction should, however, be understood against the peculiar background of humanist language studies, where Latin was always learned before Greek, and where their structural similarity helped many students to master Greek at a quicker pace than they had Latin. This situation nonetheless lays bare important weaknesses of the two modern approaches to linguistic complexity: they can result in entirely opposite conclusions and tend to presuppose that language is used, learned, and studied in a cultural vacuum. However, the case of Greek in sixteenth-century Europe suggests that the context in which the language was learned had a major impact on contemporary assessments of the language's difficulty, a fact which modern linguists can perhaps view as an invitation to factor in socio-cultural circumstances more systematically in their own assessments of the complexity of individual languages, however they interpret the concept. After all, language is not some abstract system independent of the human brain from which we can distance ourselves to objectively measure but a faculty which we can only possess, use, and study from our own personal perspectives, inevitably shaped by our cultural backgrounds.

## 5. Conclusion

The main argument I have attempted to deploy in this contribution is that sixteenth-century grammarians did not assess individual languages with a concept of complexity as it has become known in modern linguistics. Instead, the concept of *copia*, "abundance", occupied a central position in reflections on the quality of languages. *Copia* covered in part what one could consider complexity from a modern point of view; they overlapped to a certain extent<sup>21</sup>. Both *copia* and (compositional) complexity concern the number of features,

---

21. Cf. also Haßler (2009) for *copia* and related concepts, including complexity and simplicity in the seventeenth and eighteenth centuries.

forms, and categories a language has. Yet in addition to this variety, complexity also regards the interrelationships of these features, forms, and categories, which appears to have been much less the case for *copia*, if at all. This incomplete overlap might be related to the fact that language studies were not yet an autonomous discipline in the sixteenth century, whereas later notions of complexity reflected the search for patterns in language typical of modern linguistic scholarship. Indeed, *copia* was not primarily a grammatical but a rhetorical concept, whereas modern linguists have redefined complexity as a scholarly concept tailored to their own discipline, albeit with considerable variation.

In Greek-Latin contrastive observations, on which I have focused here, Greek tended to be presented as more copious overall and, somewhat contradictorily, as easy to learn, especially if one already knew Latin. The latter, in turn, was often characterized as poor and deficient, a view reflecting the linguistic prestige contest raging in the sixteenth century (Giard 1992). Unsurprisingly, eager appraisals of Greek *copia* occurred in grammars of this language, which can be expected to promote their topic. Therefore, my analysis is no doubt partly skewed because of the contextual bias of Philhellenism underlying the many positive statements on the Greek language and its richness. Remarkably enough, however, Greek *copia* was also applauded in Latin grammars, albeit often – but not exclusively – in those written by scholars with a Hellenist background, such as Melanchthon. Difficulty or ease of learning was another notion which humanist grammarians often involved in describing and assessing languages, which comes close to one modern interpretation of linguistic complexity. By this standard, Greek was surprisingly often judged to be an easy language to learn; this assessment probably served to stress the feasibility of mastering this language in an age where the relevance of Greek studies was under constant pressure (cf. Goldhill 2002: Chapter 1 for an engaging account focusing on Erasmus), very much as it is today. As such, this characterization in terms of ease of learning was also a substantiation of the ongoing linguistic battles of the time. I have cautiously argued that the relevance of context for such assessments can be taken by linguists as an encouragement to consider cultural factors in researching

complexity. Factoring in such language-external circumstances more consciously will prevent a linguist's objective complexity assessments from becoming a layperson's subjective language attitudes.

Further research, focusing on languages other than Latin and Greek, is required to find out whether my argument that sixteenth-century grammarians mainly worked with the *copia* concept rather than a complexity-like notion can be upheld. It would also be worthwhile to investigate more closely what place *copia* held in Renaissance grammars and in which other ways it was operationalized in linguistic descriptions. *Copia* was of course not the only quality attributed to languages in the sixteenth century. A somewhat contradictory quality, that of brevity, was used to applaud languages such as Hebrew and German<sup>22</sup>. Julius Caesar Scaliger, too, can be placed in the pro-brevity camp, as he promoted a rational approach to grammar, which implied conciseness and precluded the variation that was so dear to humanist stylistics. Follow-up investigations will no doubt shed light on the ways in which all these different qualifications related to each other on the linguistic battlefield of the sixteenth century.

## References

### Primary sources

Albertus, Laurentius. 1573. *Teutsch Grammatick, oder Sprach-Kunst: Certissima ratio discendæ, augendæ, ornandæ, propagandæ, conseruandæ[ue] linguæ Alemanorum siue Germanorum, grammaticis regulis et exemplis comprehensa & conscripta*. Augustæ Vindelicorum [Augsburg]: excudebat Michaël Manger.

---

22. For Hebrew, see the contribution by Maxime Maleux in this volume. For German, see e.g. Albertus (1573, sig. H 6<sup>v</sup>-H 7<sup>r</sup>): "*Breuitas & facilitas nostratis linguæ tantum unam admittit coniugationem. Quanta verò hîc sit copia prolixitas & perplexitas Latinorum, Græcorum & Hebræorum omnes docti norunt, omnesq[ue] studiosi summè deplorant, & maxima cum indignatione ac nausea ediscere coguntur.*" Cf. Jansen (1995).

- Antinori, Carlo, Urbano Bolzanio, Guarino Favorino, Aldus Manutius & Angelo Poliziano, eds. 1496. *Θησαυρός. Κέρως ἀμαλθείας, καὶ κῆποι Ἀδώνιδος. Thesaurus Cornu copiae. & Horti Adonidis.* Venetiis [Venice]: in domo Aldi Romani.
- Cauchie, Antoine. 1576. *Grammatica Gallica in III. lib. distributa.* Antverpiæ [Antwerp]: Apud Lucam Bellerum / Excudebat Ægidius Radæus.
- Ceporinus, Jacobus. 1522. *Compendium Græcæ grammaticæ.* Basileæ [Basel]: apud Valentinu[m] Curionem.
- Erasmus, Desiderius. 1534. *De duplici copia Verborum ac Rerum Commentarij duo multa accessione, nouisq[ue] formulis locupletati.* Basileae [Basel]: in officina Frobeniana.
- Gretser, Jakob. 1593. *Institutionum de octo partibus orationis, syntaxi et prosodia Græcorum, Libri Tres.* Ingolstadii [Ingolstadt]: Excudebat David Sartorius.
- Keil, Heinrich, ed. 1855-1880. *Grammatici Latini.* 7 vols. Lipsiae [Leipzig]: in aedibus B. G. Teubneri.
- Lascaris, Constantine & Aldus Manutius. 1512. *In hoc libro haec habentur: Constantini Lascaris Byzantini de octo partibus orationis Lib. I. Eiusdem de constructione liber secundus. Eiusdem de nomine et uerbo liber tertius [...].* Venetiis [Venice]: apud Aldum.
- Lopad, Ludwig. 1536. *Græcæ linguæ rudimenta, præter hactenus ædita singulari & diligentia & breuitate in puerorum usum conscripta.* Basileæ [Basel]: [Thomas Platter & Balthasar Lasius].
- Melanchthon, Philipp. 1520. *Integrae Græcæ grammatices institutiones [...] pluribus in locis auctae.* Haganœ [Haguenaue]: in ædibus Thomæ Anshelmi Badensis.
- Melanchthon, Philipp. 1526. *Elementa Latinae grammatices.* Noremburgæ [Nuremberg]: apud Io. Petreium.
- Mosellanus, Petrus. 1518. *Oratio de variarum linguarum cognitione paranda.* Lipsiae [Leipzig]: in officina Valentini Schumann.
- Perotti, Niccolò. 1489. *[Cornucopia seu commentarii linguae Latinae].* Venetiis [Venice]: per magistrum Paganinum de Paganinis Brixiansem.
- Petri, Suffridus. 1566. *Orationes quinque de multiplici utilitate linguae Graecae: inseruntur autem, praeter caetera, de optima studiorum ratione, et imitatione, quaedam, et alia lectu dignissima.* Basileae [Basel]: per Ioannem Oporinum.
- [Restauld de Caligny, Alain]. 1533. *Institutiones in Hebræam linguam.* Parisiis [Paris]: Excudebat Christianus Wechelus sibi, & Hieronymo Gormontio.
- Ruland, Martin. 1556. *De lingua Graeca eiusque dialectis omnibus libri V.* Tiguri [Zürich]: apud Andream Gesnerum F. et Iacobum Gesnerum, fratres.

- Scaliger, Julius Caesar. 1540. *De causis linguae Latinae libri tredecim*. Lugduni [Lyon]: apud Seb. Gryphium.
- Scaliger, Julius Caesar. 2019. *De causis linguae Latinae. Des causes de la langue latine*. Edition et traduction par Pierre Lardet, Geneviève Clerico et Bernard Colombat. 2 vols. Genève: Droz (Travaux d'Humanisme et Renaissance, 594).
- Sifanus, Laurentius. 1564. *Orationes duæ. Altera de laudibus linguæ Græcæ & Isocratis. Habita Coloniae Agrippinæ, Anno 1558. Pridie Nonas Iulij. Altera in commendationem Historiæ, ibidem habita Anno 1564. XIII. Kal. Maij. Coloniae* [Cologne]: Apud hæredes Arnoldi Birckmanni.
- Sylvius, Jacobus. 1531. *In linguam Gallicam Isagoge, vnà cum eiusdem Grammatica Latino-gallica, ex Hebræis, Græcis, & Latinis authoribus*. Parisiis [Paris]: ex officina Roberti Stephani.
- Winter von Andernach, Johann. 1527. *Syntaxis Graeca nunc recens, et nata, et aedita*. Lutetiae [Paris]: apud Aegidium Gormontium.

## Secondary literature

- Ax, Wolfram. 2011. *Quintilians Grammatik (Inst. orat. 1,4-8): Text, Übersetzung und Kommentar*. Berlin & Boston: De Gruyter (Texte und Kommentare, 37).
- Botley, Paul. 2010. *Learning Greek in Western Europe, 1396-1529: Grammars, Lexica, and Classroom Texts*. Philadelphia: American Philosophical Society (Transactions of the American Philosophical Society Held at Philadelphia for Promoting Useful Knowledge, 100.2).
- Burke, Peter. 2004. *Languages and Communities in Early Modern Europe*. Cambridge: Cambridge University Press.
- Callahan, Virginia W. 1978. The *De copia*: The bounteous horn. *Essays on the Work of Erasmus*, ed. by Richard L. DeMolen. New Haven: Yale University Press. 99–109.
- Denecker, Tim & Pierre Swiggers. 2018. The *Articulus* according to Latin grammarians up to the Early Middle Ages: the complex interplay of tradition and innovation in grammatical doctrine. *Glotta* 94. 127–152.
- Diu, Isabelle & Alexandre Vanautgaerden. 2018. Le jardin d'abondance d'Érasme : Le *De copia* et la lettre sur les *Adages* non éditée par P.S. Allen. *La varietas à la Renaissance*, ed. by Dominique de Courcelles. Paris: Publications de l'École nationale des chartes (Études et rencontres). 43–55.
- Fögen, Thorsten. 2000. *Patrii sermonis egestas. Einstellungen lateinischer Autoren zu ihrer Muttersprache: Ein Beitrag zum Sprachbewusstsein in der römischen Antike*. München & Leipzig: K. G. Saur.

- Giard, Luce. 1992. L'entrée en lice des vernaculaires. *Histoire des idées linguistiques*, ed. by Sylvain Auroux, vol. 2. Liège: Mardaga (Philosophie et langage). 206–225.
- Goldhill, Simon. 2002. *Who Needs Greek? Contests in the Cultural History of Hellenism*. Cambridge: Cambridge University Press.
- Haßler, Gerda. 2009. Reichtum. *Lexikon sprachtheoretischer Grundbegriffe des 17. und 18. Jahrhunderts*, ed. by Gerda Haßler & Cordula Neis, vol. 1. Berlin & New York: Walter de Gruyter. 930–945.
- Jansen, Jeroen. 1995. *Brevitas: Beschouwingen over de beknoptheid van vorm en stijl in de renaissance*, 2 vols. Hilversum: Verloren.
- Joseph, John E. & Frederick J. Newmeyer. 2012. “All languages are equally complex”: the rise and fall of a consensus. *Historiographia Linguistica* 39(2–3). 341–368.
- Lallot, Jean. 2013. Aorist (aóristos), Ancient Theories of. *Encyclopedia of Ancient Greek Language and Linguistics*, ed. by Georgios K. Giannakis. Online edition. Brill. [[https://referenceworks.brillonline.com/entries/encyclopedia-of-ancient-greek-language-and-linguistics/\\*-SIM\\_00000416](https://referenceworks.brillonline.com/entries/encyclopedia-of-ancient-greek-language-and-linguistics/*-SIM_00000416), accessed June 29, 2023].
- Mack, Peter. 2011. *A History of Renaissance Rhetoric 1380-1620*. Oxford: Oxford University Press.
- Matthaios, Stephanos & Raf Van Rooy. 2021. The Greeks' idle instrument: the article from Ancient Greece to the Renaissance. *Beiträge zur Geschichte der Sprachwissenschaft* 31(1). 27–54.
- Miestamo, Matti, Kaius Sinnemäki & Fred Karlsson. 2008. *Language Complexity: Typology, Contact, Change*. Amsterdam & Philadelphia: John Benjamins (Studies in Language Companion Series, 94).
- Newmeyer, Frederick J. & Laurel B. Preston, eds. 2014. *Measuring Grammatical Complexity*. Oxford & New York: Oxford University Press.
- Percival, W. Keith. 1983. Grammar and rhetoric in the Renaissance. *Renaissance Eloquence: Studies in the Theory and Practice of Renaissance Rhetoric*, ed. by James J. Murphy. Berkeley, Los Angeles & London: University of California Press. 303–330.
- Pontani, Anna. 2002. Jean Christophe Saladin, *La bataille du grec à la Renaissance*. Paris, Les Belles Lettres, 2000 (Histoire. Collection dirigée par M. Desgranges et P. Vidal-Naquet). Un vol. di pp. 546. *Aevum* 76(3). 852–867.
- Rescher, Nicholas. 1998. *Complexity: A Philosophical Overview*. New Brunswick & London: Transaction.
- Saladin, Jean-Christophe. 2000. *La bataille du grec à la Renaissance*. Paris: Les belles lettres (Histoire, 47). [Reprints in 2004 & 2013]

- Stevens, Benjamin. 2006–2007. Aeolism: Latin as a dialect of Greek. *The Classical Journal* 102(2). 115–144.
- Trovato, Paolo. 1984. “Dialetto” e sinonimi (“idioma”, “proprietà”, “lingua”) nella terminologia linguistica quattro- e cinquecentesca (con un’appendice sulla tradizione a stampa dei trattatelli dialettologici bizantini). *Rivista di letteratura italiana* 2. 205–236.
- Van Rooy, Raf. 2020a. *Greece’s Labyrinth of Language: A Study in the Early Modern Discovery of Dialect Diversity*. Berlin: Language Science Press (History and Philosophy of the Language Sciences, 2).
- Van Rooy, Raf. 2020b. *Language or Dialect? The History of a Conceptual Pair*. Oxford: Oxford University Press.
- Van Rooy, Raf. 2021. An ablative for the Greeks? A grammar dispute in Tübingen (1585/1586) and its implications. *Historiographia Linguistica* 47(2). 137–168.
- Van Rooy, Raf. Forthcoming. Sous l’égide de la normalisation: L’appropriation romaine de la grammaire grecque en contexte. *Le contexte en question*, ed. by Gerda Haßler. London: ISTE.



# À propos de la supériorité de l'arabe sur le persan dans un traité de grammaire arabe du x<sup>e</sup> siècle : apologie de la langue et expérience cognitive

**Julien Sibileau**

Centre de Recherches Moyen-Orient et Méditerranée (INALCO-CERMOM)

---

## RÉSUMÉ

Cet article prend pour objet la façon dont un grammairien de la TGA<sup>1</sup>, Ibn Ğinnī (m. 392<sup>2</sup>/1002), a argumenté en faveur de la supériorité de l'arabe classique sur le persan, à travers un passage de son ouvrage principal : les *Ḥaṣā'is*. Nous étudierons point par point les étapes de son raisonnement en mettant en avant sa spécificité par rapport à celui d'un savant de la même époque : Ibn Fāris (m. 395/1004), qui a lui aussi abordé le même sujet. Deux questions nous animeront tout au long de notre étude. Tout d'abord, on se demandera comment l'histoire et la politique façonnent les représentations développées par les hommes sur le langage. Et dans le même temps, nous nous demanderons si, de ce monde préscientifique, n'émergeraient pas des idées intéressantes dont les savoirs sur le langage peuvent aujourd'hui tirer profit.

## MOTS-CLÉS

hiérarchie des langues, Tradition Grammaticale Arabe, Ibn Ğinnī, supériorité

- 
1. Tradition grammaticale arabe.
  2. La première date est celle du calendrier hégirien, la seconde celle du calendrier grégorien.

rité de la langue arabe, persan, apologie de la langue arabe, langue seconde

#### ABSTRACT

This paper focuses on a grammarian of the Arabic tradition, Ibn Ğinnī (d. 392/1002), and his argument in favour of the superiority of classical Arabic over Persian developed in his main work, *al-Ḥaṣāʾiṣ*. We will discuss his reasoning step by step bringing to light its specificity as compared to the discourse of his contemporary Ibn Fāris (d. 395/1004) who tackled the same issue. Our analysis is twofold: first, we will consider the question of how history and politics shape the representations of language developed by mediaeval scholars; and second, we will try to evaluate if linguistic ideas of this prescientific period could not benefit our modern linguistic sciences .

#### KEYWORDS

language hierarchy, Arabic Grammatical Tradition, Ibn Ğinnī, superiority of Arabic, Persian, apology of Arabic, second language

## 1. Introduction<sup>3</sup>

Dans le domaine de la TGA, le discours apologétique de la supériorité de l'arabe sur le persan apparaît explicitement vers le IV<sup>e</sup>/X<sup>e</sup> siècle et repose sur le raisonnement suivant : le Coran vante les mérites de l'arabe, donc l'arabe est supérieur aux autres langues. Cette thématique qui prend racine dans un contexte où les forces de l'empire abbasside s'amenuisent, voit naître un discours hiérarchisant. Dans une configuration sociale et ethnique en plein bouleversement, où des dynasties non-arabes comme celle des Buwayhides s'imposent dans les marges de l'empire abbasside arabe, un mouvement de relecture du patrimoine littéraire arabe voit le jour<sup>4</sup>. Le traité à partir

3. Je tiens tout particulièrement à remercier Jean-Patrick Guillaume pour l'échange constructif que nous avons eu suite à ma communication dans le cadre de la conférence organisée par la SHESL sur la simplicité et la complexité des langues dans l'histoire des théories linguistiques en janvier 2020. Je tiens aussi à remercier chaleureusement Georgine Ayoub pour la relecture attentive de cet article.

4. Dans le même temps, il ne faut pas négliger le fait que les grammairiens, en affirmant la supériorité de l'arabe sur les autres langues, supériorité qu'ils prétendaient par ailleurs être les seuls à montrer, cherchaient certainement à

duquel nous allons mener la présente étude, les *Ḥaṣā'iṣ* d'Ibn Ğinnī<sup>5</sup>, s'inscrit pleinement dans ce mouvement. Dédié à un souverain buwayhite, Bahā' al-Dawla (m. 402/1012), cette pièce importante de la TGA aborde en plusieurs passages la thématique de la supériorité de l'arabe sur le persan, présentée comme une réaction face au mouvement des Persans<sup>6</sup> qui vantent les mérites de leur langue et de leur civilisation. Or, nous y reviendrons, les Buwayhides, pourtant persans, patronnaient principalement des œuvres littéraires en langue arabe.

Dans cette étude, nous voudrions prendre pour objet le discours tenu par IĜ sur ce sujet en étudiant ses arguments et en mesurant leur portée apodictique. En parallèle, nous opèrerons de la même manière avec un autre grammairien, ou plutôt lexicographe, de la TGA, Ibn Fāris (m. 395/1004), qui est l'auteur d'un ouvrage intitulé *al-Šāhibī fī Fiqh al-Luġa*<sup>7</sup>, dans lequel on trouve aussi une réflexion à propos de la supériorité de l'arabe classique sur le persan, et dont nous interrogerons les fondements. Cependant l'enjeu principal de cet article est surtout d'expliquer la tension dont le discours d'IĜ témoigne, entre discours apologétique et empirisme, en le ramenant au contexte social et politique de cette fin du x<sup>e</sup> siècle, et en interrogeant la portée de ses arguments à tendance plus « rationnelle » au regard de la linguistique. Il s'agira pour nous de mieux identifier les frontières entre un discours marqué par la mobilisation d'arguments théologiques et la rationalité grammaticale, reposant quant à elle sur des arguments d'une autre nature. Pour cela, nous organiserons notre étude en quatre parties. Nous prendrons d'abord pour objet la thématique de la supériorité de l'arabe telle qu'elle apparaît dans le discours d'IF et d'IĜ. Nous nous pencherons ensuite sur le rapport entre la notion de supériorité et ce que nous pourrions appeler « l'imaginaire grammatical ». Dans la troisième partie de cette étude,

---

asseoir leur pouvoir sur la langue face aux desseins hégémoniques, ou perçus comme tels, de la *falsafa*.

5. Désormais *Ḥ* et IĜ.

6. Dans ce travail « persan » renverra à ce qui est relatif à l'Iran et au monde iranisé islamique.

7. Désormais *Ş* et IF. On notera par ailleurs que le traité d'IF est dédié à un vizir, homme de lettres, et mécène buwayhite, al-Šāhib b. 'Abbād (m. 385/995).

nous explorerons le rapport entre l'expérience de la langue et la hiérarchie des langues et nous clorons sur le rapport entre supériorité et potentiel rhétorique de l'arabe classique.

## 2. Le thème de la supériorité de l'arabe classique

### 2. 1. La supériorité de l'arabe et le Coran

Dans le  $\text{Ş}$ , un chapitre est exclusivement consacré au thème de la supériorité de la langue arabe sur les autres langues. Ce chapitre, bien connu des chercheurs, s'intitule *Bāb al-qawl fī anna luġat al-ʿarab afḍal al-luġāt wa-awsaʿuhā* (« Que la langue des Arabes est la meilleure des langues et la plus riche »)<sup>8</sup>. De ce titre de chapitre trois informations importantes sont à retenir :

1. La notion de supériorité est exprimée à travers le superlatif *afḍal* (racine *f-ḍ-l*) qui exprime l'idée générale de vertu, de prédominance, d'excellence.
2. La notion de supériorité porte sur toutes les langues sans exception (*al-luġāt* est le *nomen rectum* du syntagme génitif dont le superlatif est la tête).
3. La notion de richesse, de latitude, de largesse, est rendue grâce à l'emploi de l'élatif *awsaʿ* (plus riche), substantif *saʿa*. Nous y reviendrons plus loin.

Toutefois la thématique du caractère distingué de l'arabe est beaucoup plus ancienne. Ainsi, Georgine Ayoub rappelle que, dès le VIII<sup>e</sup> siècle, une tradition rapportée par Ibn Sallām al-Ġumaḥī (m. 232/846) dans l'introduction des *Ṭabaqāt Fuḥūl al-Šuʿarāʾ*, et attribuée au grammairien Yūnus b. Ḥabīb (m. 182/798), affirme que « Le premier qui ait parlé la *ʿarabiyya* et qui ait oublié la langue de son père est Ismaël fils d'Abraham » (Ayoub 2017 : 80). Autrement dit la question de la langue est très tôt, dans la culture arabe, une question linguistico-religieuse (Ayoub 2017 : 81). C'est parce qu'elle est la langue des prophètes qu'elle jouit d'un statut de langue de

---

8. Seule une partie de ce passage sera exploitée ici dans le but de mieux faire ressortir la spécificité du mode de raisonnement d'IĠ.

prestige. Cette idée réapparaît un peu plus tard dans la *Risāla* du juriste al-Šāfi'ī (m. 204/823), au sein de laquelle un discours de type hiérarchisant est explicitement attesté :

*Wa-lisān al 'arab awsa' al-alsina maḡhaban wa-aktaruhā alfāzan wa-lā na'lamuhu yuḡītu bi-ḡami' 'ilmihī insān ḡayr nabī.*

La langue arabe est la plus riche, celle qui contient le plus de mots, et nous ne connaissons personne capable de la connaître toute entière si ce n'est un prophète. (Šāf., *Risāla*, 42)

On note donc ici que l'expression *awsa' al-alsina* sera reprise environ deux siècles plus tard par IF dans son *Ṣ* (*awsa'uhā*) et que le caractère distingué de l'arabe est associé au fait que seul un prophète peut l'embrasser dans sa totalité. Autrement dit, la supériorité de l'arabe va de pair avec son statut de langue sacrée. Au début du x<sup>e</sup> siècle, le philosophe Abū Ḥātim al Rāzī (m. 322/934) consacre un chapitre à la supériorité de l'arabe sur les autres langues intitulé *Faḍl luḡat al-'Arab* (« La supériorité de la langue des Arabes ») dans son *Kitāb al-Zīna*, où il affirme dans un premier temps que quatre langues, parmi lesquelles l'arabe, sont supérieures aux autres : l'arabe, l'hébreu, le syriaque, le persan, puis que parmi ces quatre langues, c'est l'arabe qui se distingue (*inna afdal al-luḡāt al-arba' luḡat al-'arab*). Cette distinction se manifeste ici à travers l'expression *afdal al-luḡāt* et il suffit de lire le reste du chapitre pour se rendre compte qu'elle s'articule aussi avec le domaine du sacré, notamment avec l'argument du caractère intraduisible du Coran (*wa-lā qadara aḡad min al-umam an yutarḡimahu bi-ṣay' min al-alsina*)<sup>9</sup>. À la même époque, cette thématique est intégrée au programme de la grammaire. Dans l'introduction du *Kitāb al-Uṣūl* son auteur, le grammairien Ibn al-Sarrāḡ (m. 316/929), évoque un niveau de causalité qui permettrait d'appréhender la sagesse (*ḡikma*) déposée dans les principes (*uṣūl*) de la langue arabe et de mettre en évidence sa précellence sur toute autre langue (*faḍl ḡāḍihī l-luḡa 'alā ḡayrihā min al-luḡāt*)<sup>10</sup>. Par conséquent, pour lui, la supériorité de la langue arabe trouverait sa justification dans l'organisation même de son système qui refléterait la sagesse de son fondateur (Guillaume 2012 : 286-287). En somme, on

9. Voir Rāzī., *Zīna*, 73.

10. Voir Sarr., *Uṣ.*, t. 1, 35. Voir aussi Zaḡḡ., *Īd.*, 64-66.

peut donc dire que le thème de la supériorité de l'arabe est attesté dès le VIII<sup>e</sup> siècle, qu'il est dû au caractère sacré de la langue, mais revêt de multiples formes dans les ouvrages de la culture arabe classique à mesure que celle-ci se développe. Aussi, compte tenu de l'ampleur de la question, nous avons choisi de restreindre notre étude à deux discours, l'un attribué à IĠ et l'autre à IF, en nous focalisant sur la nature de leurs arguments visant à défendre l'idée de la supériorité de l'arabe sur les autres langues. Cette brève mais néanmoins nécessaire mise au point étant faite, revenons au texte d'IF.

La prémisse sur laquelle le lexicographe fait reposer son discours apologétique est tirée d'un passage du Coran s'étendant sur quatre versets 192-195 de la sourate « Les poètes » dans lequel la langue arabe est qualifiée de *mubīn* (« claire »). La racine de ce participe actif du verbe de forme IV *abāna b-y-n* apparaît par ailleurs dans un autre dérivé, celui de *bayān*, littéralement l'exposition claire, l'intelligibilité. Voici le verset en question :

*Wa-innahu la-tanzīl-u rabb-i l-‘ālamīna* (192) *Nazala bi-hi l-rūḥ-u l-amīn-u ‘alā qalbi-ka* (193) *Li-takūn-a min al-munḍirīna* (194) *Bi-lisān-in ‘arabiyy-in **mubīn-in*** (195) (Coran 26, 192-195)

Et il est une révélation du seigneur des mondes (192) Descendue par l'esprit fidèle sur ton cœur (193) afin que tu sois parmi les avertisseurs (194) Dans une langue arabe claire (195).

Les commentateurs ont suggéré maintes interprétations de ce verset. C'est le raisonnement hiérarchisant établi par notre lexicographe à partir de ce verset qui nous intéresse ici. Pour lui, l'acte énonciatif d'une révélation en langue « claire », implique *ipso facto* l'idée que l'arabe classique doit occuper le sommet de la hiérarchie des langues<sup>11</sup> :

*Fa-lammā ḥaṣṣa ḡalla ṭanā'uhu al-lisān al-‘arabī bi-l-bayān ‘ulima anna sār al-luḡāt qāsira ‘anhu wa-wāqī‘a dūnahu.*

Puisque Dieu a spécifiquement mentionné la langue arabe et l'a qualifiée de claire, il faut en déduire que les autres langues sont dépourvues de cette qualité et se trouvent donc être inférieures<sup>12</sup> à l'arabe. (IF, §, 19)

11. Le texte en question d'IF a été traduit par Georgine Ayoub (2006).

12. Ici l'adverbe *dūna*, c'est-à-dire « en dessous », « inférieur », exprime clairement l'idée de hiérarchie.

Lorsqu'on rétorque à IF que la clarté est un attribut dont l'ensemble des langues se parent, celui-ci s'enferme dans un dogmatisme. Sa réponse est néanmoins intéressante d'un point de vue linguistique en ce que les principaux concepts du système communicationnel développé par les linguistes y sont attestés :

*Kull man afhama bi-kalāmihi 'alā šarṭ luġatihi fa-qad bayyana.*

Quiconque se fait comprendre par l'intermédiaire de son discours selon le code de sa langue a produit un discours clair. (IF, §, 19)

Le *bayān* est donc associé à la présence d'un locuteur dont la volonté est de se faire comprendre par l'intermédiaire de son discours selon le code de sa langue. « Se faire comprendre » est rendu en arabe par le verbe factitif de forme IV *afhama*, dérivé du verbe de forme I *fahima yafhamu fahm* (« comprendre »). Le discours quant à lui, est exprimé à l'aide du terme *kalām*, terme polysémique s'il en est, qui désigne dès le *Kitāb* de Sībawayhi à la fois le discours et la langue (Ayoub 2010 : 26-28). En revanche, dans l'extrait qui nous occupe ici, le concept de langue est bien rendu en arabe par un terme spécifique, celui de *luġa*, qui désigne ici la langue en tant que système. En outre le syntagme prépositionnel *'alā šarṭ luġatihi* semble indiquer l'idée de code<sup>13</sup>. Il n'y a de compréhension (*fahm*) qu'en raison de la présence d'un code connu à la fois par l'émetteur du message et le récepteur, et c'est ce processus qui permet au locuteur de rendre son discours clair.

Cette remarque relève bien évidemment du sens commun. Mais le lexicographe s'entête. Pour lui, dire que les locuteurs peuvent produire du sens et se faire comprendre grâce à des codes différents est une lapalissade. La hiérarchie<sup>14</sup> des langues doit s'opérer selon des critères rhétoriques. Ainsi, au plus bas de l'échelle de l'exposition se trouve le mode d'expression basique, celui consistant à se faire comprendre de son interlocuteur de la manière la plus simple qui soit. Quant au *bayān*, il constituerait un mode d'expression dont le fonctionnement obéirait à des principes plus complexes que le discours basique :

13. Voir Ayoub (2006 : 44).

14. Souvenons-nous du *dūna-hu* mentionné plus haut dans le propos d'Ibn Fāris.

*In kunta turīd anna l-mutakallim bi-ğayr al-luğa l-‘arabiyya qad yu‘rib ‘an nafsihi ḥattā yafham al-sāmi‘ murādahu fa-hādā aḥass marātib al-bayān li-anna l-abkam qad yadull bi-išārāt wa-ḥarakāt lahu bi-akṭar murādihi ṭumma lā yusammā mutakallim faḍlan ‘an an yusammā bayyin wa-baliğ.*

Si tu veux dire que le locuteur peut s’exprimer à travers une autre langue que la langue arabe de manière à ce que le récepteur comprenne ce qu’il a voulu dire, cela représente le moyen d’expression le plus bas, en ce que le muet peut tout à fait indiquer une grande partie de ce qu’il veut dire grâce à des signes et des gestes, ce qui ne fait pourtant pas de lui un locuteur, encore moins un expositeur clair ou un rhéteur. (IF, §, 19)

De nouveau, il appert que le discours d’IF tourne essentiellement autour de la notion de hiérarchie, une notion rendue en arabe par le terme de *marātib*. Pluriel de *martaba*, il renvoie à l’idée d’organisation d’objets, concrets et abstraits, selon un ordre déterminé. En l’occurrence, il s’agit dans ce contexte des modes d’expression. En haut de la pyramide, nous l’avons compris, se situe l’arabe classique avec, selon IF, son haut potentiel rhétorique. En bas de l’échelle, se trouve la langue des signes du muet, indicative, c’est le verbe *dalla yadullu* (« indiquer ») qui est employé. Dans ce passage, chaque attribut possède une signification précise, le *mutakallim*, simple locuteur, n’est pas le *bayyin*, locuteur mobilisant des stratégies plus complexes afin de mettre en forme ses idées (*ma‘ānī*), ni un *baliğ*, locuteur s’attachant quant à lui à adapter son discours en fonction de son interlocuteur. Le lexicographe est formel : pour lui, les langues ne possèdent pas toutes le même potentiel rhétorique et, afin de défendre l’idée selon laquelle celui de l’arabe serait supérieur aux autres langues, il invoque toute une série d’arguments<sup>15</sup> :

1. La synonymie (*al-tarāduf*) *id est* la richesse lexicale de l’arabe
2. Les figures de rhétorique
  - 2.1. La métaphore (*al-isti‘āra*)
  - 2.2. La comparaison (*al-tamṭīl*)
  - 2.3. L’inversion (*al-qalb*)
  - 2.4. Le sous-entendu (*al-taqdīr*)
  - 2.5. La postposition (*al-ta’ḥīr*)

---

15. Ces arguments ne seront pas étudiés dans ce travail mais pourront faire l’objet d’une étude ultérieure.

Pour IF, la supériorité de l'arabe s'incarne donc à travers la richesse de son potentiel rhétorique bien que ces figures soient attestées dans la plupart des traditions rhétoriques du monde. En outre, le lexicographe précise bien que sa réflexion ne porte que sur la langue du Coran (*sunan al-‘arab fī l-Qur’ān*)<sup>16</sup>. Toutefois, derrière ce discours idéologique, une idée précise semble se dessiner, pour nous qui étudions les représentations des langues développées par les hommes au cours de l'histoire, celle voulant que la supériorité d'une langue dépende de son potentiel rhétorique. D'après cette théorie médiévale, plus le potentiel rhétorique d'une langue serait élevé, plus la langue en question serait « supérieure ». Si le scientifique voulait explorer la question du potentiel rhétorique des langues, il lui faudrait adopter une méthode contrastive et porter un jugement objectif sur toutes les langues de son corpus. À l'époque médiévale, dans la culture arabe, l'idée dominante est que la perfection de l'arabe classique tient à son statut de langue sacrée, comme nous l'avons vu plus haut. À partir d'un argument scripturaire, IF pense avoir montré que l'arabe classique jouit effectivement d'un statut linguistique beaucoup plus élevé que les autres langues. On ne manquera pas de noter sa mauvaise foi puisque le verset coranique n'exprime pas le moins du monde l'idée de supériorité à laquelle il fait pourtant allusion. Par ailleurs la TGA abonde en représentations hiérarchisantes sur le langage. L'une d'entre elles se démarque tout particulièrement des autres par la nature des arguments qu'elle mobilise.

## 2. 2. La supériorité de l'arabe et le locuteur

Si IĠ et IF ont en commun de faire reposer leur discours apologétique sur des arguments rhétoriques, il n'en demeure pas moins que dans les *H*, l'argument scripturaire vu chez IF ne figure point, même dans l'un des chapitres les plus polémiques que celui-ci ait écrit : *Bāb fī anna l-‘Arab qad arādat min al-‘ilal wa-l-aġrād mā nasabnāhu ilayhā wa-ḥamalnāhu ‘alayhā* (« Chapitre à propos du fait que les Arabes ont bien voulu des causes et des finalités que nous leur avons attribuées »). La problématique essentielle que tente de traiter

---

16. IF, §, 19.

IĠ dans ce chapitre tourne autour du thème de la correspondance entre les explications des grammairiens, la théorie et la réalité des faits de langage. Cependant, son raisonnement fait état d'un vide épistémique et non des moindres car, à son époque, les Bédouins puristes, ceux s'exprimant de manière spontanée en conservant les désinences finales, ont presque tous disparu. Comment accéder, dès lors, à la « réalité » de la langue ? Comment vérifier la portée aléthique des énoncés explicatifs produits par les grammairiens ? Pour combler ce manque, IĠ mobilise des récits dont le personnage central est un bédouin, 'Abdallāh al-Šaġarī, appartenant à la prestigieuse tribu des Banū Tamīm, plus précisément au clan des 'Uqayl, avec lequel il se serait lié d'amitié à Mossoul. Cette stratégie fait vraisemblablement écho aux enquêtes « linguistiques » d'antan où les grammairiens fréquentaient encore les Bédouins dans le but de recueillir des données. Ibn Ġinnī est indéniablement un grammairien nostalgique, épris d'un retour vers les fondements de la grammaire, d'une redécouverte des principes cachés de cette discipline.

Dans une de ces anecdotes, on voit le grammairien tester les compétences du bédouin dans le domaine de la morphologie nominale. En principe, la règle stipule que les noms communs en  $C_1u/iC_2C_3ān$  aient un pluriel interne en  $C_1aC_2āC_3īn$  comme *dukkān* (« boutique ») > *dakākīn* ou encore *sirhān* (« loup ») > *sarāhīn*<sup>17</sup>. Mais comme toutes les règles, celle-ci possède son exception. Si le nom en question est un nom propre, son pluriel est un pluriel externe, c'est-à-dire qu'il convient de suffixer au prénom le morphème *-ūna* des pluriels humains comme dans *'Utmān* (Osman) > *'Utmān-ūna* (IĠ, *H.*, t. 1, 257). Émerveillé par la réussite du bédouin à ce test, le grammairien y voit la preuve que les règles conçues par ses pairs sont bien réelles, situées dans la psychologie des locuteurs, pratiquées par ces derniers de manière inconsciente. On ne comprend pourtant pas très bien le rapport entre cette anecdote et le thème de la supériorité. A priori, cette anecdote ne nous dit rien si ce n'est que le langage du bédouin obéit à des lois dont ce dernier est inconscient. Toutefois,

---

17. Nous ne nous attardons pas ici sur la question des critères d'interprétation des gabarits. La racine de *dukkān* peut tout à fait être interprétée comme *d-k-k* ou *d-k-n*. Nous nous en tenons à l'interprétation d' IĠ car ce thème ne fait pas partie de l'objet de cet article.

celle-ci s'achève sur une remarque du bédouin sur laquelle va reposer ce que nous appellerons l'« imaginaire grammatical » :

*A-ra'ayta insān yatakallam bi-mā laysa fī kalāmihi.*

As-tu déjà vu un homme employer des choses étrangères à sa langue ?  
(IĠ, *H*, t. 1, 257-258)

Bien que le contenu paraisse relativement neutre, le bédouin ne faisant qu'exprimer un principe universel que tout locuteur peut ressentir lorsqu'il s'agit de prononcer une expression non conforme aux règles de sa langue, IĠ ne manque pas de l'interpréter comme la marque de l'amour passionné qu'éprouveraient les locuteurs à l'encontre de leur langue (*šagāfihim bi-luġatihim*), de leur vénération envers elle (*ta'zīmihim lahā*), et de leur croyance ferme qu'elle serait la plus belle des merveilles (*wa-'tiqādihim aġmal al-ġamīl fihā*)<sup>18</sup>. Le décalage entre la neutralité de la réaction du bédouin et le discours dithyrambique du grammairien étonne. Mais pour le second, si l'arabe classique mérite des qualificatifs aussi prestigieux, c'est en raison de leur conservatisme. Par conséquent, bien qu'IĠ tienne un discours apologétique à l'égard de l'arabe, tout comme son homologue, il n'en demeure pas moins qu'il se réfère d'abord au locuteur.

### 3. Conservatisme et imaginaire grammatical

#### 3. 1. La question de l'emprunt

Pour notre grammairien, le fait que le système de l'arabe classique soit si conservateur relèverait, non pas d'une propriété intrinsèque au système, mais de la volonté des locuteurs de ne pas intégrer d'éléments étrangers à leur langue. Pour lui, cette ferme volonté de repli sur soi de la communauté linguistique constituerait une preuve irréfutable de la supériorité de ses membres, en ce qu'elle indiquerait leur volonté de préserver leur langue tel un joyau que rien ne saurait altérer. Dans le chapitre 34 de notre traité, celui qui nous occupe ici, maintes stratégies sont mises en œuvre par IĠ dans le but d'illustrer cette idée. C'est le cas par exemple d'un vers de poésie attribué à Abū

---

18. IĠ, *H*, t. 1, 258.

Mahdiyya, poète de la tribu de Bāhila, dans lequel celui-ci exprime son vif rejet des mots empruntés au persan :

*Yaqūlūna l-ī šanbiḍ wa-lastu mušanbiḍ-an \* tiwāl-a l-layālī mā aqāma tabīr-u*

*Wa-lā qā'il-an zūd-an li-ya'ǧal-a šāhib-ī \* wa-bistān fī šadrī 'alayya kabīr-u*

*Wa-lā tārik-an lahnī li-uḥsin-a lahn-a-hum \* wa-law dāra šarf-u l-dahr-i ḥaytu yadūr-u*<sup>19</sup>

Ils me disent « *šanbiḍ* » mais je ne les saluerai point en persan \* tant que le monde sera monde et que *Tabīr*<sup>20</sup> s'élèvera dans les cieux.

Tout comme je ne dirai point « *zūd* » pour que mon compagnon se dépêche \* ni « *bistān* », quelle ignominie !

Je n'abandonnerai pas mon dialecte pour le leur \* quelles que soient les épreuves qui se dresseront en travers de mon chemin<sup>21</sup>. (IĠ, *H*, t. 1, 255)

Dans un premier temps, il est important d'attirer l'attention sur le fait qu'environ deux siècles et demi séparent ces vers de la période qui nous occupe ici, ce qui montre que le discours d'IĠ est tiraillé entre le fait que la langue dont ce dernier vante les mérites est celle d'un passé révolu, et la volonté, comme nous venons de le dire, de faire comme si cette dernière était encore parlée par quelques Bédouins puristes pour des raisons vraisemblablement liées à la légitimation de la grammaire arabe<sup>22</sup>. Cette relecture montre aussi un aspect de la dimension littéraire des *H* où l'auteur mobilise et arrange des anecdotes anciennes en vue d'orienter son lectorat vers une idée : celle de la perfection de la langue arabe. Par ailleurs, ce court poème nous intéresse tant sur le plan linguistique que sur celui de l'histoire des représentations épilinguistiques car son auteur exprime avec ardeur sa volonté de ne pas employer des mots

19. Les astérisques marquent la séparation entre deux hémistiches.

20. Montagne dans le voisinage de La Mecque (IM, *Lisān*, p. 470).

21. On pourra aussi se référer à la traduction des mêmes vers dans Ayoub (2006 : 42-43).

22. Dans la culture arabe classique le témoignage oral constitue une source d'informations reconnue. Dès lors, on comprend qu'IĠ veuille se targuer, à l'instar des fondateurs de la grammaire, d'avoir été en contact direct avec les locuteurs de la *'arabiyya*, langue dont il fait l'apologie.

arabes empruntés au persan<sup>23</sup>. Le verbe *šanbada*, qui n'est attesté dans aucun dictionnaire d'arabe classique, que ce soit le *Lisān* ou le *Kazimirski*, se compose de trois consonnes *š-b-d* qui correspondent probablement aux trois consonnes de l'expression persane *če būd-ī* (litt. « comment étais-tu ? »), à savoir l'affriqué *č*, la labiale *b* et la dentale *d*, la liquide *n* aurait été incrémentée en position de C<sub>2</sub> comme c'est le cas d'un bon nombre de néologismes arabes, qu'ils soient d'ailleurs empruntés ou non. Cette vue est confirmée par le lexicographe arabe al-Ġawālīqī (m. 540/1144) qui glose *šanbid* par *šūn būdī* dans son *Mu'arrab* (Ġaw., *Mu'a.*, 98), et que l'éditeur rectifie en *dġūn būdī* (« comment étais-tu ? » > « comment vas-tu ? »), ce qui correspond à notre intuition de départ. Quant à *zūd*, il provient aussi du persan *zūd* qui signifie littéralement « têt » en persan moderne. Enfin *bistān* provient lui aussi du persan. C'est l'impératif du verbe *setadan* (« prendre ») à la 2<sup>ème</sup> personne du singulier. Par ailleurs, il est intéressant de noter la performance métalinguistique de notre poète qui, au premier hémistiche du second vers, produit un discours sur la langue, puisqu'il attribue à *zūd* une fonction conative en disant que celui-ci sert à presser le récepteur du message.

Ce genre d'anecdote est utile d'un point de vue linguistique mais aussi sociolinguistique, malgré la rareté des indices renfermés par cette dernière. En effet, celle-ci témoigne du comportement linguistique d'un membre de la tribu de Bāhila qui, nous le savons, pérégrinait au centre de la péninsule arabe (Rabin 1951 : 14). Or on sait combien le parler de ces tribus de l'intérieur jouissait d'un statut de prestige aux yeux des philologues arabes : *afṣaḥ al-'arab abarruhum* (Rabin 1951 : 18). Cependant on peut tout à fait se demander si ce conservatisme ne serait pas lié aux exigences explicites ou implicites de la poésie. S'agit-il d'une tendance générale, attestée dans le discours ordinaire du poète ainsi que dans ses productions poétiques, ou bien celle-ci se limite-t-elle au domaine de la poésie ? S'il est difficile de répondre à cette question, il en est une autre dont la réponse semble davantage à notre portée, celle concernant l'identité des destinataires du poème. À qui fait référence le pronom suffixe de la troisième personne *-hum* dans *laḥn-a-hum* ? Le terme de *laḥn*, nous le savons, est polysémique,

23. Pour davantage de détails sur cette question, on pourra se référer à Ayoub 2006 : 57-62.

puisqu'il peut désigner à la fois la manière de parler, la mélodie et le chant, ainsi que la faute de langage (Ayoub 2001 : 71-73). Ici, le terme renvoie incontestablement à une façon de parler. Deux communautés linguistiques employaient le persan dans le cadre de leurs activités quotidiennes : les Persans, dont le persan était bien évidemment la langue maternelle, et les Arabes situés aux confins de l'ancien empire sassanide, comme ce fut le cas des Lakhmides, où le contact géographique entre les deux populations a certainement dû engendrer des transferts entre les deux systèmes linguistiques. Cependant, les trois termes dont il est question dans le poème sont arabes, empruntés au persan certes, mais ils ont été, pour reprendre un terme de (Weinreich 1968 : 44), « grammaticalement intégrés ». Autrement dit, l'auteur de ces vers s'adresse indéniablement à des Arabes dont le parler mêle des éléments de persan, un comportement linguistique qu'il fustige violemment.

La présence du verbe de forme IV *aḥsana* dans *li-uḥsina laḥna-hum* mérite aussi quelques commentaires. Celui-ci possède une valeur positive puisqu'il désigne le fait de bien effectuer quelque chose. Opposé à *taraka*, « délaissé, abandonner », il marque la volonté du poète de ne pas adopter la manière de parler d'un autre groupe, qu'il considère comme davantage exposé au contact linguistique, et semble donc indiquer une opposition entre parlars de l'intérieur et ceux de la périphérie<sup>24</sup>. Bien évidemment, cet argument en faveur de la supériorité de l'arabe sur le persan ne possède aucune valeur apodictique puisque, comme le signale à juste titre l'interlocuteur d'IĠ, les Persans sont eux aussi passionnés par leur langue (*bi-luġatihim mašġūfūna*), la préfèrent à d'autres (*lahā mu'tirūna*), répugnent aussi à l'emprunt de mots arabes (*wa-an yadhūlahā šay' min al-ʿarabiyya kārihūna*) et le poète qui compose en persan s'expose à de vives critiques s'il emploie des mots arabes (*iḏā awrada l-šā'ir minhum šī'r fihī alfāz min al-ʿarabī ʿiba bīhi wa tu'ina li-aġli dālika ʿalayhi*)<sup>25</sup>.

Il ne faut certainement pas prendre cette remarque à la légère car il est fort probable que celle-ci fasse écho à l'entreprise littéraire

24. D'après Ingham en effet, le système des dialectes de l'intérieur (*dialects of inner Arabia*) seraient davantage marqués par rapport à celui des dialectes de la périphérie (*outer dialects*). Voir Ingham 1982 : 33.

25. IĠ, *H*, t. 1, 258.

menée par le fameux poète Abū l-Qāsem Ferdowsī. Né en 329/940 à Tūs, l'actuel Mašhad, cet homme de lettres est l'auteur d'un poème épique faisant peu ou prou office de la *Chanson de Roland* en France, le *Šāh Nāme* (« Livre des Rois »). Or cette compilation des légendes persanes dans laquelle son auteur narre l'histoire de l'Iran depuis la création du monde jusqu'à l'arrivée de l'Islam, et dont la composition s'étend d'environ 366-67/966-967 à 400/1010, s'inscrit dans un processus de conservation et de défense du persan<sup>26</sup> face à l'invasion de l'arabe, comme l'indique lui-même l'auteur. Cette œuvre majeure de la littérature mondiale en langue persane, patronnée au départ par la dynastie des Samanides<sup>27</sup>, IĠ ne pouvait l'ignorer<sup>28</sup>. Ceci est d'autant plus vraisemblable que le poète persan avait ses entrées à la cour de Bahā' al-Dawla<sup>29</sup>. Autrement dit, cette discussion ne semble pas être apparue à cet endroit, à ce moment-là de l'histoire, dans ce traité de grammaire arabe, de manière fortuite. Elle semble plutôt s'inscrire dans le cadre d'une rivalité intellectuelle indéniablement

---

26. La question de savoir si Ferdowsī était versé dans la connaissance du pehlevi fait encore l'objet de désaccords parmi les chercheurs.

27. Nous savons en effet que l'œuvre de Ferdowsī s'inscrit dans la continuité de celle de Daqīqī, qui répond par ailleurs à la volonté du souverain samanide de l'époque, Abū Šāliḥ Mansūr ou Manšūr Ier (961-976). Pour plus de détails sur la vie de Ferdowsī, voir l'article que lui consacre Khaleghi-Motlagh (1999) dans l'*Encyclopædia Iranica*.

28. On pourra à cette occasion mentionner le vers du *Šāh Nāme* dont l'attribution à Ferdowsī est contestée :

*Basī ranġ bordam badīn sāl sī \* Aġam zande kardam badīn (be īn) pārsī*

Thirty long years I laboured [on this construction] \* I revived ancient Persia ('Ajam) in this Persian tongue (*be-dīn Pārsī*)

Ce vers est mentionné en note de bas de page par Djalal Khaleghi-Motlagh dans son édition critique du *Šāh Nāme* (Khaleghi-Motlagh 1997, t. 8 : 487). La traduction est d'Alireza Shapur Shahbazi (Shapur Shahbazi 1991 :123-124). Que ce vers soit de Ferdowsī ou non, il n'en reste pas moins que le *Šāh Nāme* représente une œuvre majeure de la littérature persane, sinon son œuvre fondatrice.

29. D'après Claude Cahen, les Buwayhides ont fait montre d'un intérêt, bien que mineur, pour la littérature néo-persane. Il affirme néanmoins que leur rôle dans ce domaine n'a pas eu l'importance de celui des Samanides. Voir Cahen (2010 : 296) ainsi que l'article de l'*Encyclopédie de l'Islam* qu'il consacre aux Buwayhides (Cahen 2007 [1960]).

très enracinée, et dont les réflexions sur le langage de l'époque témoignent.

Si les Persans sont aussi fiers de leur langue que le sont les Arabes de la leur, que reste-t-il donc comme argument en faveur de la supériorité de l'arabe ? On a par ailleurs le sentiment qu'IĠ confond deux phénomènes distincts. La première anecdote dans laquelle son transmetteur, al-Šaġarī, refuse de prononcer un mot afin de respecter la règle grammaticale, n'a rien à voir avec le phénomène de diglossie dont il est question dans le poème d'Abū Maḥdiyya. Toute cette confusion nous amène à établir le constat suivant : c'est l'imaginaire grammatical, ce mythe de l'intuition linguistique bédouine, qui l'emporte très souvent face à l'observation objective des faits.

### 3. 2. Supériorité et égalité : le lexique de la hiérarchie

Dans cette rêverie qui emporte la raison d'IĠ, un lexique de la supériorité et de l'égalité émerge. C'est ainsi que l'interlocuteur d'IĠ s'interroge : « Les deux langues sont donc égales à ce niveau, en quoi l'arabe serait-il supérieur au persan ?<sup>30</sup> » (*fa-qad tasāwat hāl al-luġatayni fī dālīka fa-ayyatu faḍīla li-l-'arabiyya 'alā l-'aġamiyya*). Si l'émerveillement des membres d'une communauté linguistique envers leur langue, ainsi que le conservatisme de ces derniers dans leurs comportements linguistiques, semblaient être l'argument ultime permettant de prouver la supériorité de cette langue par rapport aux autres, le fait que l'on retrouve ces deux traits dans d'autres cultures constitue un contre-argument fort. À ce stade, rien, dans le discours d'IĠ, ne permet de pencher pour la supériorité de l'arabe sur le persan, ou plutôt, aucun argument sérieux n'est avancé. De fait, à ce niveau-là de la lecture de notre corpus, la notion d'égalité prime, et elle est rendue en arabe à travers le verbe de forme VI *tasāwā* (« être de même niveau »), dont la racine est *s-w-y*, et que l'on retrouvera

---

30. Sur la question des noms des langues dans la TGA, voir Guillaume (2009 : 64) où il explique que les noms des langues étrangères comme le grec ou le persan, sont « simplement conçues dans leur relation, plus ou moins vague, avec un groupe et/ou un territoire ». C'est précisément le cas pour *'aġamiyya*, adjectif de relation construit à partir de *'aġam*, terme qui désigne au départ les non-arabes, synonyme de βάρβαροι, puis finit par être quasi-synonyme de *Furs*, *i.e.* Perses.

un peu plus loin dans le même passage. Cette égalité s'applique entre deux langues, en témoigne le suffixe du duel *-ayni* dans *luġat-ayni*. Quant à la notion de supériorité, elle est exprimée à travers l'emploi d'un dérivé de la racine *f-d-l* construit sur le schème C1aC2iC3a > *faḍīla* (« mérite, qualité supérieure »).

Cette position intellectuelle repose sur deux critères que notre texte exprime à travers l'emploi d'un démonstratif (*fī ḍālika*, littéralement « en cela »), que nous venons de mentionner : l'attachement des locuteurs à leur langue et leur conservatisme. Si le premier ne possède de valeur apodictique que celle que l'imaginaire du grammairien lui attribue, le second peut suggérer de façon très subliminale, l'idée d'un rapport entre complexité, dans son acception structurelle, et isolation. Si on abstrait le discours d'IĜ de son contexte idéologique, on peut dire que l'idée générale qui en ressort est celle d'un rapport entre pureté, dans le sens de préservation des reliquats et autres archaïsmes, et isolation. Plus une langue serait isolée, moins son système serait susceptible d'être altéré, et donc d'être simplifié<sup>31</sup>.

Cependant il faut bien reconnaître que dans notre exemple la complexité semble se trouver du côté de la périphérie et du contact. En effet, si l'on admet que la complexité se trouve toujours du côté de la multiplicité (règles de grammaire, flexions, etc.), il nous faut donc prendre en compte le phénomène de la synonymie. Plus un signifié posséderait de signifiants, plus le système linguistique en question serait complexe. Or les phénomènes de contact contribuent grandement, on le sait, à enrichir le vocabulaire d'une langue, parfois même à produire des formes linguistiques hybrides. De ce point de vue, c'est-à-dire au niveau terminologique, la complexité se trouverait du côté de la périphérie et du contact. Quoi qu'il en soit, il est important de noter que les grammairiens ne se préoccupent absolument pas des effets du contact, sauf IĜ à de rares occasions dans les *H*. La discussion, chez les grammairiens de l'arabe, tourne seulement autour du contact, et on lui attribue la plupart du temps

---

31. Une idée d'ailleurs énoncée par Rafaella Baechler en ces termes : « it is expected that languages spoken by language communities with many contacts tend to simplify their language system whereas languages spoken by isolated language communities (i.e with few contacts) tend to preserve complexity as to complexify their language system. » (Baechler 2017 : 179).

une valeur négative, d'où la paire *an-yadḥulahā / kārihūna*, c'est-à-dire « contact = rejet » puisque la pureté (*faṣāḥa*) de la langue en dépend. IĜ reste néanmoins le grammairien qui a probablement le plus discoursé sur ces deux aspects du contact : l'aspect idéologique, c'est-à-dire le contact en tant que phénomène doté d'une valeur négative, et l'aspect linguistique, c'est-à-dire les effets du contact sur le système de la langue.

## 4. Hiérarchie des langues et expérience cognitive

### 4. 1. L'acquisition d'une langue seconde

À ce niveau de l'analyse, on mesure bien l'écart entre le raisonnement d'IF et celui d'IĜ. Le premier, nous l'avons vu, faisait reposer sa démonstration sur un argument d'autorité qu'il avait détourné de son sens. Le verset coranique constituait le point de départ de sa réflexion. Du côté d'IĜ, il en va tout autrement car à aucun moment celui-ci ne puise dans le texte coranique afin d'étayer son raisonnement. Peut-on expliquer cet écart en revenant aux deux grands courants qui dominaient la sphère intellectuelle au IV<sup>e</sup>/X<sup>e</sup> siècle, à savoir le *naql*, méthode basée sur la transmission orale, et le *'aql*, méthode quant à elle reposant sur la raison et l'expérience ? À la lecture du discours de nos grammairiens, nul doute que cette dichotomie épistémologique se manifeste bien et semble avoir joué un rôle important dans le développement de cette discussion. Chez IĜ, la méthode empirique ne se tapit pas à l'ombre de son discours, elle est explicite. En témoigne le nombre élevé d'attestations de dérivés de la racine *ḥ-s-s*. Dans le corps du traité, on n'en compte pas moins de 27 occurrences, dont le verbe de forme IV *aḥassa bi-*, « sentir quelque chose ». Pour IĜ, la question de la supériorité (*faḍīla*) s'articule à celle de la sensation, et si les Persans, affirme-t-il, avaient « senti » la subtilité de l'art des locuteurs (*ṣinā'at al-'arab*) dans cette langue, son caractère mystérieux (*ġumūd*), sa précision (*diqqa*) ainsi que sa finesse (*riqqa*), ils auraient alors cessé de vanter les mérites de leur langue. Cette fois-ci, c'est le substantif verbal *taqdīm* qui est employé afin d'exprimer l'idée de prévalence d'une chose sur une

autre. *Qaddama* A 'alā B signifie « faire prévaloir A sur B », où la préposition locative remplit à nouveau le rôle de *mapping* logique entre A et B.

Ce discours est ensuite neutralisé par l'interlocuteur d'IĠ, qui rétorque que les Persans pensent la même chose de leur langue. Surgit alors un argument empirique faisant intervenir le témoignage de grammairiens de l'arabe d'origine persane, c'est-à-dire dont la langue maternelle est le persan, et l'arabe la langue seconde. Justement affirme IĠ, lorsqu'on interroge les grammairiens de l'arabe ('ulamā' al-'arabiyya) d'origine persane (*mimman ašluhu 'aġamī*), qui ont étudié leur langue maternelle avant d'apprendre l'arabe, ces derniers sont unanimes : ils affirment que les deux langues ne sont pas égales (*lā yaġma' baynahumā*)<sup>32</sup>. Dans ce passage, deux termes ont trait au champ lexical de la cognition, il s'agit d'abord du verbe de forme V *tadarraba* qui signifie « s'exercer », « s'entraîner à » avec la préposition 'alā, ainsi que du verbe intransitif de forme X *ista'raba*, qui exprime ici l'idée d'apprentissage de la langue arabe classique. Quant aux formes et aux modalités de cet apprentissage, aucun détail n'est donné.

Le témoignage d'Abū 'Alī l-Fārisī, affirme son fidèle disciple, renforce cette conviction, et les témoignages de grammairiens de l'arabe d'origine persane, poursuit-il, affluent aussi en ce sens. Selon lui, les compétences acquises par ces derniers en arabe leur permettraient de renforcer leur connaissance du persan, de mieux connaître cette langue, et d'être plus attentif à ses états (*fa-inna quwāhum fī l-'arabiyya tu'ayyid ma'rifatahum bi-l-'aġamiyya wa-tu'annisuhum bihā wa-tazīd fī tanbīhihim 'alā aḥwālihā*)<sup>33</sup>. La réflexion d'IĠ ne repose donc pas sur un argument d'autorité relevant du religieux, comme c'était le cas chez IF, mais sur la conviction profonde que l'acquisition de l'arabe en tant que langue seconde a permis aux grammairiens non-arabophones de développer la connaissance de leur langue maternelle. Dans un langage abstrait, dépouillé de toute la rêverie qui entoure cette assertion, on pourrait dire que c'est l'apprentissage d'une langue seconde, en général, qui

32. IĠ, *H*, t. 1, 258.

33. IĠ, *H*, t. 1, 258.

peut contribuer au développement de la connaissance de la langue maternelle chez les sujets parlants. En somme, il aurait suffi qu'IG remplace « arabe » par « langue seconde » pour que nous ayons là les prémisses d'une réflexion linguistique de portée générale. Pourtant, dans l'ensemble des textes de la Tradition, IG est peut-être le grammairien ayant le plus contribué à l'abstraction du raisonnement grammatical. De ce discours apologétique dont le texte des *H* est indiscutablement imprégné, surgit parfois de lumineuses réflexions qui n'ont pas leur pareil dans toute la Tradition.

#### 4. 2. Une grammaire universelle ?

La réflexion à laquelle nous venons de faire écho à propos de l'apport cognitif lié à l'acquisition de l'arabe par un persanophone, s'accompagne en outre d'une brève réflexion sur les grammaires et leurs structures qui, loin d'offrir des réponses définitives à nos questions, ouvre des pistes de recherche tout à fait stimulantes concernant d'éventuelles influences étrangères dans les *H*. En général, nous le savons, les grammaires se font discrètes quant à la méthodologie qu'elles renferment. Rares sont les discours épistémologiques proposés par les grammairiens sur les concepts grammaticaux et le raisonnement grammatical en général. Là encore, il semble que les *H* occupent une place tout à fait particulière dans l'histoire de cette tradition. Quoi qu'il en soit, l'apport cognitif dont nous venons de parler ne constitue pas un simple exergue mais un processus mental structuré dont les principaux traits sont les suivants. Il y a d'abord ce qu'IG appelle les compétences des Persans en arabe, qu'il rend en arabe par le pluriel de *quwwa* > *quwā*, littéralement « la force ». Ces compétences engendrent ensuite trois conséquences :

1. Renforcement de la connaissance du persan (*tu'ayyid ma'rifatahum bi-l-ʿağamiyya*)
2. Familiarisation avec cette langue (*tu'annisuhum bihā*)
3. Attention plus élevée envers ses états (*tazīd fī tanbīhihim ʿalā aḥwālihā*)

Pour nous, ce genre de séquence, qui semble basée sur la redondance et la répétition, cache des nuances sémantiques subtiles qu'il nous incombe de décrypter. Cette séquence, qui se décompose

en trois temps, débute par la mention de la connaissance (*ma'rifa*). Ce terme est si général qu'il ne nous est malheureusement pas possible de spéculer à son propos. Il peut s'agir de connaissances en prose, en poésie, dans le domaine de la syntaxe, de la morphologie, de la phonologie. Toutes les hypothèses sont envisageables. En revanche, le champ sémantique du second niveau semble plus réduit puisqu'il y est question du verbe de forme II *'annasa*, « rendre familier », de la racine *'n-s*, dont les mots *'uns* (« intimité, familiarité ») et *'ins* (« genre humain ») dérivent. L'emploi de cette racine, selon toute vraisemblance, indique donc l'idée de familiarité dans l'usage de la langue car un individu ne peut s'accoutumer à une langue que s'il en maîtrise les différents usages. Enfin, l'association du verbe de forme II *nabbaha*, « éveiller, avertir », avec le pluriel de *ḥāl* > *aḥwāl*, possède selon nous une connotation beaucoup plus technique. En effet, à de très nombreuses reprises dans les *H*, ce terme s'articule avec des réflexions phonétiques, morphologiques et syntaxiques. Comme nous le montrons ailleurs (Sibileau 2021 : 217-226), il correspond à un mode ontologique.

Mais ce qu'il y a de plus intéressant encore, c'est l'explication de ce phénomène apportée par IĠ. Pour lui, cet apport cognitif n'est possible qu'en vertu de l'interconnexion des savoirs linguistiques entre eux (*ištirāk al 'ulūm al luġawiyya wa štibākihā*) et du but commun à leurs significations (*wa tarāmihā ilā l-ġāya l-ġāmi'a li-ma'ānīhā*)<sup>34</sup>. Cette dernière expression, quelque peu mystérieuse, revêt un intérêt tout particulier en ce qu'il y est question des *ma'ānī l-'ulūm al-luġawiyya*, littéralement les « significations des savoirs linguistiques ». Afin d'éclaircir ce passage, il suffit de revenir sur les multiples occurrences de ce que certains grammairiens, parmi eux IĠ, appellent les *ma'ānī l-i'rāb*<sup>35</sup>, et que les spécialistes de la TGA appellent la sémantique grammaticale. Si l'on se fie au texte des *H*, le *ma'nā* désigne en général un intelligible pouvant par définition être « saisi par l'âme » (*ma'nā ta'qiluhu l-nafs*)<sup>36</sup>. Ces *ma'ānī* touchent de nombreux domaines de la langue, principalement ceux de la

34. IĠ, *H*, t. 1, 258

35. L'*i'rāb* renvoie probablement ici à la grammaire (*naḥw*) de manière générale.

36. IĠ, *H*, t. 1, 192.

syntaxe et de la morphosyntaxe. Chacune des parties du discours est qualifiée de *ma'nā*. Il est question de la signification verbale (*ma'nā l-fi'liyya*)<sup>37</sup>, de la signification nominale (*ma'nā l-ismiyya*)<sup>38</sup> et enfin de la signification prépositionnelle (*ma'nā l-ḥarfīyya*)<sup>39</sup>. Deux des quatre cas de l'arabe, le nominatif (*raf*) et le génitif (*ḡarr*), au moins, sont qualifiés de *ma'ānī*, ce qui nous permet de dire que les cas en général relèvent de la sémantique grammaticale<sup>40</sup>. Le duel (*taṭniya*) est aussi un *ma'nā*<sup>41</sup>, de même que la négation (*naḡy*)<sup>42</sup>. Enfin, la définition (*ta'rīf*) et l'indéfinition (*tankīr*) sont aussi qualifiées de *ma'nā*<sup>43</sup>. De fait, lorsqu'IĠ évoque les « significations » des savoirs linguistiques, on est en droit de penser que celui-ci fait allusion au fait que certaines langues, qu'il ne nomme pas, disposeraient de catégories plus ou moins communes. Mais que désigne alors le terme de *ḡāya* ?

Si notre lecture de ce passage est correcte, IĠ ferait ici l'exposition de sa vaste culture linguistique, dépassant le cadre de l'arabe. Au sein de l'aire géographique dans laquelle il a vécu, étaient parlés le grec, le syriaque, le persan, et l'arabe. Nous savons que la tradition grammaticale syriaque débute au VI<sup>e</sup> siècle avec la traduction de la *Technè grammatikè* du (Pseudo) Denys le Thrace attribuée au syriaque oriental Joseph Huzaya et que la première grammaire du syriaque proprement dite, le *Tūraṣ Memlla* de Jacques d'Édesse, apparaît dans la seconde moitié du VII<sup>e</sup> siècle (Farina 2021). Quant à la première grammaire du persan, celle-ci n'apparaît que bien tardivement, au XIX<sup>e</sup> siècle, avec le *Mollā Bāšī* de 'Abd al-Karīm Iravāni (Jeremiás 1993 : 52). Cependant, on peut présumer, d'après ce que nous avons dit plus haut, que les réflexions sur cette langue remontent à beaucoup plus loin. Ceci étant dit, le concept de *ḡāya* auquel IĠ fait allusion pourrait renvoyer à la finalité de la grammaire qui vise à ordonner les

---

37. IĠ, *H*, t. 1, 157.

38. IĠ, *H*, t. 1, 532.

39. IĠ, *H*, t. 1, 534.

40. IĠ, *H*, t. 1, 290-291.

41. IĠ, *H*, t. 2, 191.

42. IĠ, *H*, t. 1, 322.

43. IĠ, *H*, t. 1, 343.

faits linguistiques de la manière la plus économique possible de façon à faciliter l'apprentissage d'une langue. Or cet ordonnancement des faits se caractérise par l'identification de catégories abstraites dans lesquelles seront rangés les faits linguistiques, et ces catégories, bien qu'elles ne soient pas toutes identifiables dans les différentes langues<sup>44</sup>, n'en sont pas moins nécessaires à la conception d'une grammaire. Que serait une grammaire qui ne s'intéresserait pas au nombre, à la définitude, à la flexion, etc. ? Si les savoirs linguistiques convergent tous vers un but commun, c'est tout simplement parce qu'ils reposent sur un socle de *ma'ānī*, des catégories abstraites permettant de systématiser l'agencement des faits linguistiques. Nous pouvons aller plus loin dans la spéculation et faire le lien avec l'apport cognitif dont il était question plus haut. Si l'apprentissage d'une langue seconde par un sujet contribue à mieux connaître sa langue maternelle, c'est peut-être parce que le caractère commun des catégories grammaticales lui permet de comparer le système des deux langues.

Une branche de la linguistique prend pour objet l'acquisition des langues secondes. Pour Rod Ellis, reprenant les travaux de Larry Selinker (Selinker 1972), le sujet apprenant une seconde langue construit une série de grammaires mentales, ou d'interlangues, au fur et à mesure que sa connaissance de la L2 augmente (Ellis 1997 : 33). Cette grammaire mentale se construit à partir de diverses stratégies employées par le sujet parlant, comme le transfert de la L1 ou encore l'hypercorrection (Ellis 1997 : 34 ; VanPatten & Williams 2007 : 39). Autrement dit le sujet apprenant est amené à raisonner à un niveau abstrait où les catégories grammaticales et les règles sont sans cesse ajustées par rapport à la structure grammaticale de la L1 (Keating 2016 : 25), ce qui lui permet de changer de point de vue par rapport à cette dernière. En fait, moins L2 possède de traits en commun avec L1, plus le coût d'acquisition de L2 est élevé pour l'apprenant puisqu'il n'est pas en mesure de raisonner en termes de transfert de L1 vers L2. Ainsi, un locuteur de l'anglais aura plus de difficultés à appréhender l'ordre non-SVO de l'espagnol qu'un locuteur de l'italien (VanPatten & Williams 2007 : 122). Toutefois, le fait que L2

---

44. Par exemple, le persan ne connaît ni la distinction de genre ni la définitude.

contienne des traits non attestés en L1 n'implique en rien que celle-ci soit plus complexe, ou dans la langue d'IĠ, supérieure. Quoi qu'il en soit, IĠ réaffirme un peu plus loin que les savants d'origine persane comme le lecteur coranique Abū Ḥātim al-Siġistānī (m. 248/862), le traditionniste Bundār (m. 252/865) ainsi que son maître, ne situent pas les deux langues au même niveau (*lam nara aḥad min ašyāḥinā fihā ka-Abī Ḥātim wa-Bundār wa-Abī 'Alī wa-fulān wa-fulān yusawwūna baynahumā wa-lā yuqarribūna bayna ḥālayhimā*)<sup>45</sup>. Nous retrouvons donc ici, à nouveau, la racine *s-w-y* exprimant la notion d'égalité à travers le verbe de forme II *sawwā*, « aplanir, niveler, mettre deux choses ou plus sur le même plan ». En somme, comme nous venons de le voir, c'est l'approche rationaliste d'IĠ qui le distingue, et de loin, d'un savant plus dogmatique comme l'est IF.

## 5. Le potentiel rhétorique de l'arabe

### 5. 1. Flexibilité (*sa'a*) et poésie

Bien que le mode de réflexion des deux savants soit aux antipodes, un point commun ressort malgré tout de l'étude de leur discours, celui du rapport entre supériorité de la langue et rhétorique. Si la discussion d'IĠ à propos de la supériorité de l'arabe sur le persan a touché à sa fin, maints passages du traité semblent faire résonner certains des termes employés par IF. Parmi ceux-ci, celui de *sa'a*. Certes, c'est le superlatif *awsa'* qui est attesté dans le *Ṣ*, mais ce terme s'articule malgré tout avec les figures de la rhétorique. Qu'en est-il dans les *H*? L'impression dont nous a fait part IĠ dans le chapitre 34 reste lettre muette car aucune perspective comparatiste n'est réellement envisagée. Pour lui, la théorie n'a qu'une seule vocation : montrer la dimension rationnelle, harmonieuse, et parfaite de la langue arabe. Mais cette perfection, qui se traduit dans son langage à travers le terme arabe de *ḥikma* qui littéralement signifie « sagesse » (Bohas 1981), s'articule souvent avec le terme de *sa'a* (littéralement « espace, latitude, largesse »), auquel Kees Versteegh a d'ailleurs consacré un article important (Versteegh 1990), où il explique que ce terme renferme l'idée de liberté du locuteur, en particulier dans le domaine

45. IĠ, *H*, t. 1, 258-259.

de la syntaxe<sup>46</sup>. Comment cette notion s'incarne-t-elle concrètement dans le raisonnement d'IĠ ?

Dans le chapitre 160 de notre traité, intitulé *Bāb fī l-ġam' bayna l-ad'af wa-l-aqwā fī 'aqd wāhid* (« Chapitre à propos de l'emploi d'une expression faible et d'une expression forte dans un même énoncé »), il est question d'énoncés un peu particuliers qui mêlent à la fois une expression « faible », c'est-à-dire s'écartant du *qiyās* (« analogie »), et une expression « forte », c'est-à-dire conforme à l'analogie et employée fréquemment. Pour illustrer ce phénomène, IĠ cite le vers d'un célèbre poète arabe du VIII<sup>e</sup> siècle, al-Farazdaq :

*Kilā-humā hīna ġadda l-ġary-u bayna-humā \* qad aqlā'ā wa kilā anfay-himā rābī.*

Tous deux, alors que la course s'intensifiait \* **étaient partis** depuis belle lurette et **chacun** de leurs nez **s'élevait au ciel**.<sup>47</sup>

Dans ce vers, le quantifieur *kilā* est dans un premier temps interprété comme un duel, le verbe *aqlā'ā* en position de prédicat étant marqué du suffixe du duel *-ā*, puis, dans le second hémistiche, le prédicat de *kilā* est cette fois-ci accordé au singulier : *rābī*. De fait nous sommes en face d'un même morphème doté de deux valeurs que le locuteur, ici un poète, semble pouvoir choisir à sa guise :

1. <i>Kilā humā</i>	<i>qad</i>	<i>aqlā'ā</i>
Q.Tous Suff. Duel	Marq. énonc.	Verbe : démarrer-Acc-duel
2. <i>Kilā</i>	<i>anfay-himā</i>	<i>rābī</i>
Q. Chacun	leurs deux nez	PA : élevé-sing.

Le premier type d'accord est associé par IĠ à une valeur faible, car il dépend du référent (*li-annahū ḥaml 'alā l-ma'nā*), tandis que le second type d'accord est quant à lui considéré comme fort, car il dépend du signifiant, *kilā* n'étant pas, pour notre grammairien, marqué du duel (*li-annahū ḥaml 'alā l-lafz*)<sup>48</sup>. D'un côté, l'ensemble

46. En fait, ce concept semble très proche de celui de *copia* auquel Érasme a même consacré un ouvrage. On sait que chez lui ce concept désigne une sorte d'embellissement du discours et d'abondance du style. Je remercie d'ailleurs Raf Van Rooy ainsi que Maxime Maleux pour leurs brillants exposés ainsi que les riches échanges que nous avons pu avoir lors de la conférence.

47. IĠ, *H*, t. 2, 505.

48. IĠ, *H*, t. 2, 509.

[A1 ; A2] est interprété comme une somme d'unités, de l'autre comme un tout. Par conséquent, *kilā*, dans le premier hémistiche, peut être traduit par « tous », tandis que dans le second, il doit être traduit par « chacun ». Pour IĠ, cette flexibilité dans le domaine de l'accord en nombre du quantifieur *kilā* illustre à merveille le potentiel rhétorique de l'arabe classique, phénomène dont il précise que les locuteurs sont conscients : « Les locuteurs peuvent employer des expressions en laissant de côté d'autres expressions qu'ils préfèrent pourtant aux premières » (*qad yasta'milūna min al-kalām mā ġayruhu ātar fī nufūsihim minhu*)<sup>49</sup>. Pour lui, ce phénomène aurait pour fonction de procurer un espace de liberté discursive aux locuteurs (*sa'at-an fī l-tafassuḥ*), un relâchement du souffle (*irḥā'an li l-tanaffus*) et marquerait une volonté de conserver des pratiques langagières qu'ils ont instituées (*ṣuḥḥ-an 'alā mā ġašimūhu fa-tawāḍa'ūhu*)<sup>50</sup>. Nombreux sont les exemples de ce genre dans les *H*. La flexibilité de l'accord en nombre est interprétée comme la marque du génie des locuteurs. Or c'est justement à ce génie que l'expression *luṭf ṣinā'at al-'arab fihā* vue plus haut fait allusion. C'est cette subtilité (*luṭf*) qui conduit IĠ à affirmer la supériorité de l'arabe par rapport au persan.

## 6. Conclusion

En guise de conclusion, nous pouvons dire que le thème de la supériorité de l'arabe, bien que commun à nos deux savants, n'est pas tout à fait appréhendé de la même manière par ces derniers. Si le sensible constitue bien le point de départ de la réflexion d'un grammairien rationaliste comme IĠ, c'est le texte coranique qui prime chez IF. Si l'on observe attentivement leurs discours, on constate même que le fond diffère. La notion de *sa'a* touche essentiellement le lexique chez IF tandis qu'elle s'applique aussi à la syntaxe chez IĠ, ainsi que nous l'avons vu dans la dernière partie. Le discours du premier se limite principalement à la langue coranique, tandis que celui de son homologue porte surtout sur la poésie. La sensation occupe une place centrale dans le raisonnement de l'auteur des *H*. Elle est complètement absente chez IF. IĠ, nous l'avons vu, est

49. IĠ, *H*, t. 2, 509.

50. IĠ, *H*, t. 2, 509.

persuadé que l'acquisition de l'arabe en tant que langue seconde peut contribuer à développer les compétences des Persans dans leur langue maternelle en vertu de l'universalité de la grammaire, une intuition juste, dans le fond, mais qui n'est jamais érigée en objet d'étude. Cette étude a donc révélé une tension entre d'un côté, un discours apologétique certainement cultivé par des forces sociales et politiques, avec lesquelles IĠ était contraint de composer, et de l'autre, une volonté de sa part d'explorer les effets de l'apprentissage de l'arabe en tant que langue seconde sur la cognition des sujets persanophones.

Par ailleurs, il serait naïf de penser que cette discussion n'est pas liée à l'épistémologie de la grammaire. La flexibilité évoquée par IĠ est surtout celle de la théorie grammaticale face à la diversité des données dont les grammairiens disposaient. Ce point, qui a déjà été soulevé par Georges Bohas et & Aḥmad al-Qādirī (1998), est le principal mobile de l'évolution de la théorie grammaticale, car la démonstration de la perfection de la langue passe nécessairement par la nécessité de concevoir des outils adaptés à cette fin. En effet, ce sur quoi on se fonde est fondé à son tour sur ce qui en dépend (Meschonnic 1975). À chaque fois que la règle de grammaire n'est plus pertinente pour embrasser les données, la théorie doit se tenir prête à expliquer ce qui s'apparente à une anomalie dans le but de conserver à tout prix l'idée que le système de la langue arabe est cohérent. Autrement dit, système linguistique et théorie grammaticale vont de pair. Quel intérêt pouvait donc avoir un grammairien aussi subtil qu'IĠ à montrer la supériorité de la langue arabe ?

Il ne faut pas oublier que le traité des *H* est dédié à un souverain Buwayhīde. Or cette dynastie était farouchement opposée à celle des Samanides<sup>51</sup> qui, nous l'avons dit, promouvait une sorte de retour à la langue persane. On peut donc penser que celle-ci a cherché à prendre le contre-pied intellectuel des Samanides en favorisant une littérature qui peignait un arabe en tant que langue de prestige. Il ne faut pas s'étonner de voir deux dynasties persanes s'affronter sur

---

51. Nous savons par exemple qu'en 983 les Samanides luttèrent âprement face, entre autres, aux Buwayhīdes pour maintenir leur hégémonie dans le Ḥurāsān. Voir l'article que Bosworth et Crowe (1960) consacrent aux Samanides dans l'*Encyclopédie de l'Islam*.

un tel sujet. Aujourd'hui encore, ces débats font couler beaucoup d'encre en Iran. N'oublions pas, nous l'avons dit, qu'IĠ a grandi au sein d'une société où trois grandes civilisations cohabitaient, où seule la maîtrise de l'arabe permettait d'accéder aux hautes sphères d'une société abbasside qui se délitait. La langue arabe permettait aux minorités, dont IĠ faisait partie<sup>52</sup>, d'accéder aux hautes fonctions de l'empire, secrétariat de chancellerie, monde des Lettres, savoirs majeurs comme la théologie et la jurisprudence, philosophie. C'est donc cette langue qui lui a permis de s'élever socialement. La couvrir d'éloges, vanter ses mérites, revenait à adresser un panégyrique à l'encontre du peuple conquérant et dominateur. Enfin, on peut penser, étant donné le nombre élevé de personnifications de la langue arabe attestées dans les *H*, qu'elle symbolisait aussi pour ce grammairien fils d'esclave, une sorte de lignage qui faisait de lui un sujet légitime, dans une société où la généalogie (*nasab*) occupait un rôle si important. D'ailleurs, cette idée n'apparaîtrait-elle pas dans un de ses vers :

*Fa-in uṣbiḥ bi-lā nasabīn \* Fa 'ilmī fi l-warā nasabī.*

Bien que je vive sans généalogie \* ma science compense cette absence<sup>53</sup>.

## Bibliographie

### Sources primaires

- Ḥam., *Mu'ğam* = Al-Ḥamawī, Yāqūt. *Mu'ğam al-Udabā*, éd. par Iḥsān 'Ab-bās. Beyrouth : Dār al-ğarb al-islāmī. 1993.
- Ġaw., *Mu'a*. = Al-Ġawālīqī, Abū Maṣṣūr. *Al-Mu'arrab*, éd. par F. 'Abd al Raḥīm. Damas : Dār al-qalam. 1990.
- Rāzī, *Zīna* = Al-Rāzī, Abū Ḥātim. *Kitāb al-Zīna*, éd. par Ḥusayn b. Fayḍ Allāh al-Hamadānī. Sanaa : Markaz al-dirāsāt wa-l-buḥūṭ al-yamanī. 1994.
- Šaf., *Risāla* = Al-Šāfi'ī, Muḥammad b. Idrīs. *al-Risāla*, éd. par Aḥmad Šākir. Maṭba'at muṣṭafā : al-bābī al-ḥalabī. 1940.
- Zağğ., *Īd*. = Al-Zağğāğī, Abū l-Qāsim. *Al-Īdāḥ fi 'Ilal al-Naḥw*, éd. par Māzin al-Mubārak. Le Caire : Dār al-'urūba. 1959.

52. N'oublions pas qu'il était le fils d'un esclave byzantin.

53. Ḥam, *Mu'ğam*, p. 1586.

- Sarr., *Uṣ.* = Ibn al-Sarrāğ, Abū Bakr. *Al-Uṣūl fī l-Naḥw*, éd. par ‘Abd al-Ḥusayn al-Fatī. Beyrouth : Mu’assasat al-risāla. 1996.
- IF, *Ṣ.* = Ibn Fāris, Abū l-Ḥusayn Aḥmad. *al-Ṣāhibī fī fiqh al-luġa*, éd. par Aḥmad Ḥasan Basağ. Beyrouth : Dār al-kutub al-‘ilmiyya. 1997.
- IĠ, *Ḥ.* = Ibn Ġinnī, Abū l-Faṭḥ ‘Uṭmān. *al-Ḥaṣā’iṣ*, éd. par ‘Abd al-Ḥamīd Hindāwī. Beyrouth : Dār al-kutub al-‘ilmiyya. 2008.
- IM, *Lisān* = Ibn Manzūr, Abū l-Faḍl. *Lisān al-‘Arab*. Le Caire : Dār al-ma‘ārif. 1980.

## Sources secondaires

- Ayoub, Georgine. 2017. Langue, mythe et histoire : La ‘arabiyya, langue du paradis, des prophètes et des tribus de l’Arabie antique. *La Langue et le Sacré*, dir. par Georgine Ayoub et Alessandro Guetta. Paris : Geuthner. 73-126.
- Ayoub, Georgine. 2006. L’autre et ses mots : L’énonciateur inintelligible. *Les Mots Voyageurs et l’Orient, Contact des Langues II*, dir. par Michel Bozdémir & Sonel Bosnali. Istanbul : Presses Universitaires de Bogazici. 15-62.
- Ayoub, Georgine. 2010. De l’exemple et de l’exemplarité du vers dans le *Kitāb* de Sibawayhi. *L’exemple et la citation dans l’œuvre des grammairiens et des lexicographes arabes*, dir. par Hassan Hamzé & Bassam Baraké. Beyrouth : Dar al-Hilal. 17-91.
- Ayoub, Georgine. 2001. Le tout de la langue ou le malheur de l’infini : Une étude de la Durrat al-Ghawwās de Harīrī. *Paroles, Signes, Mythes, Mélanges J.-E Bencheikh*, dir. par Floréal, Sanagustin. Damas : Institut Français d’Études Arabes à Damas. 67-141.
- Baechler, Rafaella. 2017. Complexity, isolation, and language change. *Zeitschrift für Dialektologie und Linguistik* 84(2/3). 178-201.
- Bosworth, Clifford E. & Yolande Crowe. 2007 [1960]. Sāmānids. *Encyclopaedia of Islam 2*, dir. par Peri Bearman, Thierry Bianquis, Clifford E. Bosworth, E. J. van Donzel & Wolfhart Heinrichs. Leyde : Brill.
- Bohas, Georges. 1981. Quelques aspects de l’argumentation et de l’explication chez les grammairiens arabes. *Arabica* 28(2-3). 204-221.
- Bohas, Georges & Aḥmad al-Qādirī. 1998. Des faits, des grammairiens et des linguistes. *Arabica* 45(3). 297-319.
- Cahen, Claude. 2010 [1968]. *L’Islam : Des origines au début de l’Empire ottoman*. Paris : Fayard.
- Cahen, Claude. 2007 [1960]. Buwayhids or Būyids. *Encyclopaedia of Islam 2*, dir. par Peri Bearman, Thierry Bianquis, Clifford E. Bosworth, E. J. van Donzel & Wolfhart Heinrichs. Leyde : Brill.

- Ellis, Rod. 1997. *Second Language Acquisition*. Oxford : Oxford University Press.
- Farina, Margherita. 2021. La tradition grammaticale syriaque comme « extension » de la grammaire grecque : les parties du discours et le participe. *Histoire Épistémologie Langage* 43(1). 41-60.
- Guillaume, Jean-Patrick. 2009. Les noms des langues en arabe. *Histoire Épistémologie Langage* 31(2). 49-66.
- Guillaume, Jean-Patrick. 2012. À propos d'un fragment du *Marāḥ al-Arwāḥ*. *En mémoire de Sophie Kessler-Mesguich*, dir. par Jean Baumgarten *et al.* Paris : Presses de la Sorbonne nouvelle. 283-292.
- Ingham, Bruce. 1982. *North East Arabian Dialects*. New York & Londres : Routledge.
- Jeremiás, Éva M. 1993. Tradition and innovation in the native grammatical literature of persian. *Histoire Épistémologie Langage* 15(2). 51-68.
- Khaleghi-Motlagh, Djalal. 1999. Ferdowsī Abū l-Qāsem. *Encyclopædia Iranica*. t. 9, Fasc. 5. 514-523.
- Khaleghi-Motlagh, Djalal. 1997. *The Shahnameh*. New York : Mazda Publishers and Bibliotheca Persica.
- Keating, Gregory. 2016. *Second language acquisition: The basics*. New York & Londres : Routledge.
- Meschonnic, Henri. 1975. *Le signe et le poème*. Paris : Gallimard.
- Rabin, Chaïm. 1951. *Ancient west-arabian*. London : Taylor's Foreign Press.
- Selinker, Larry. 1972. "Interlanguage" in *International Review of Applied Linguistics* 10. 209-231.
- Shapur Shahbazi, Alireza. 1991. *Ferdowsī : A Critical Biography*. Cambridge : Harvard University Press.
- Sibileau, Julien. 2021. *Une entreprise de légitimation de la grammaire arabe : les Ḥaṣā'is d'Ibn Ğinnī (m. 392/1002)*. Thèse de Doctorat, INALCO.
- VanPatten, Bill & Jessica Williams. 2007. *Theories in Second Language Acquisition*. Mahwah (NJ) & Londres : Lawrence Erlbaum Associates.
- Versteegh, Kees. 1990. Freedom of the speaker? The term *ittisā'* and related notions in arabic grammar. *Studies in the History of Arabic Grammar II*, dir. par Kees Versteegh and Michael Carter. Amsterdam : John Benjamins. 281-293.
- Weinreich, Uriel. 1968 [1953]. *Languages in Contact: Findings and Problems*. La Haye, Paris, New York : Mouton de Gruyter.

# Les langues d'Asie, une complexité cachée

**Alice Vittrant**

Université Lumière Lyon 2/ CNRS-DDL (UMR 5596)

---

## RÉSUMÉ

Loin de certaines croyances ayant eu cours dans l'histoire de la linguistique, il règne actuellement un consensus sur l'égale complexité des langues indépendamment de leur type morphologique ou de leur affiliation. Dans un grand nombre de travaux, la complexité d'une langue est évaluée en termes de nombre et de variété des éléments (morphèmes) constituant l'item langagier. La complexité grammaticale d'une langue se mesure alors en termes de distinctions exprimées, généralement de manière obligatoire. Ainsi du point de vue de la morphologie, on associera complexité avec flexion et caractère obligatoire d'une expression.

Or les langues d'Asie du Sud-est et d'Asie orientale nous racontent une autre histoire. Elles suggèrent que cette « complexité exprimée » (*overt complexity*) n'est qu'un type particulier de complexité, et que le fonctionnement éminemment pragmatique de ces langues (interprétation basée sur le contexte) cache une complexité d'un autre genre, une « complexité cachée » (*hidden complexity*), rarement décrite qui prend sa source dans la propension des langues d'Asie à dire peu en laissant deviner beaucoup.

## MOTS-CLÉS

complexité, typologie, langues d'Asie du Sud-Est, pragmatique, contexte, polyfonctionnalité

## ABSTRACT

Far from beliefs that were held in the history of linguistics, there is currently a consensus on the equal complexity of languages, regardless of their morphological type or their affiliation. In many publications, the complexity of a language is evaluated in terms of the number and variety of elements

(morphemes) that constitute the language itself. The grammatical complexity of a language is then measured in terms of the distinctions expressed, generally in an obligatory way. Therefore, from the point of view of morphology, complexity will be associated with inflection and the *obligatoriness* of an expression.

However, the languages of Southeast Asia and East Asia tell us a different story. They suggest that this ‘overt complexity’ is only a particular type of complexity, and that the essentially pragmatic functioning of these languages (with context-based interpretation) hides a complexity of another kind, a ‘hidden complexity’, rarely described, which has at its source the propensity of Asian languages to say little while leaving a lot to guess.

#### KEYWORDS

complexity, typology, South-East Asian languages, pragmatics, context, polyfunctionality

---

## 1. Introduction

Cet article a pour but de discuter les critères utilisés pour évaluer la complexité d’une langue. L’accent sera mis ici sur un type de complexité rarement décrit, à savoir la « complexité cachée » (*hidden complexity*), qui prend sa source dans la propension des langues d’Asie à dire peu en laissant deviner beaucoup.

La première partie sera consacrée à un bref rappel du traitement de la complexité en linguistique et plus particulièrement en typologie depuis le XIX<sup>e</sup> siècle. Puis sera posée la question de l’évaluation de la complexité, celle des critères utilisés, lesquels mesurent généralement la complexité en termes de distinctions exprimées, dans la continuité des travaux du XIX<sup>e</sup> et XX<sup>e</sup> siècle. La dernière partie mettra en lumière une complexité d’un autre ordre, caractéristique de nombreuses langues d’Asie. Cette « complexité cachée » (*hidden complexity*) généralement ignorée dans la littérature jusque récemment (Bisang 2004) sera illustrée à partir de données issues des langues d’Asie du Sud-est, connues pour l’absence ou l’omission généralisée d’expressions linguistiques, la rareté des morphèmes grammaticaux obligatoires et la polyfonctionnalité importante de nombreux morphèmes (Vittrant & Watkins 2019). Ces langues peuvent être

décrites comme très pragmatiques, c'est-à-dire comme nécessitant la prise en compte systématique du contexte d'interlocution.

## 2. La complexité : un relatif consensus

### 2. 1. Historique

Il était assez courant au XIX<sup>e</sup> siècle d'associer COMPLEXITÉ avec beauté, de faire rimer SIMPLICITÉ avec capacités mentales réduites ; les philologues de l'époque assimilaient le plus souvent le caractère complexe des éléments d'une langue à sa supériorité, une supériorité dans le domaine cognitif mais aussi culturel et social comme illustré par les deux extraits suivants.

Dans leurs *Études de linguistique et d'ethnographie* (1878), Hovelacque & Vinson reprennent la classification morphologique qui avait cours à la fin du XIX<sup>e</sup> siècle, proposée par August Schlegel. Cette classification hiérarchisait et évaluait les langues sur la base de leur morphologie et donnaient bien évidemment la primauté aux langues flexionnelles indo-européennes, les langues « les plus abouties » selon la logique évolutionniste sous-jacente à cette classification :

Sous le rapport de leur structure, les langues parurent à Schlegel se grouper en trois grandes classes. La première classe, la plus simple, comprenait les idiomes sans aucune structure grammaticale. [...] Ces idiomes sont appelés isolants ou monosyllabiques. [...]

Dans la seconde classe, dans la classe des langues agglutinantes, **un progrès considérable a été opéré**. Certaines racines ont perdu une partie de leur valeur propre et primitive. [...]

De même que l'agglutination est sortie du monosyllabisme, de même la flexion sortit de l'agglutination. (Hovelacque & Vinson 1878 : 22-23)

Dans la première [catégorie] se rangent les langues de l'Asie sud-orientale [...]. Elles sont caractérisées par l'absence complète de l'expression des relations. [...] C'est là évidemment **l'état le plus imparfait du langage**. [...]

Dans la seconde [catégorie] se placent de très nombreux idiomes [...]. Ces idiomes **sont supérieurs aux précédents**, car ils savent exprimer la relation. [...] Il est donc probable qu'à l'origine les langues de ce groupe étaient isolantes et monosyllabiques. [...]

Il suit de là que **le système linguistique le plus parfait** sera celui qui indique la relation par un changement dans la forme de la racine significative qui restera une. C'est le procédé employé par les langues du troisième groupe ; c'est ce qui les distingue d'une façon bien tranchée des autres produits de l'organisme vocal humain. [...] et il est très-vraisemblable, leurs racines significatives étant du reste monosyllabiques, que toutes ces langues à flexions ont été primitivement isolantes, puis agglutinantes, et que **la flexion n'est qu'un perfectionnement ultérieur**.

De cet examen résultent aussi d'autres conséquences importantes. Si les langues à flexion sont **les mieux organisées** et si elles sont passées successivement par **deux états plus défectueux**, il faut conclure que le langage est essentiellement progressif, variable, et modifiable dans le sens d'une amélioration constante. (*Ibid.* : 55-57 ; souligné en gras par nous)

McElvenny (2017 : 14) note que le lien entre caractéristiques formelles d'une langue et évolution cognitive et sociale de ses locuteurs reste présent même chez des auteurs qui remettent en cause la classification de Humboldt postulant la supériorité des communautés linguistiques de langues flexionnelles, ce que rapporte aussi Szmrecsanyi & Kortmann (2012 : 7) :

*Wilhelm von Humboldt put forward somewhat unfortunate claims to the effect that differences (in terms of complexity or otherwise) between languages can be traced back to the differential mental capacities of their speakers.* (Szmrecsanyi & Kortmann 2012 : 7)

Le point de vue anthropologique d'une equi-richesse des cultures et des logiques de pensée reste une position minoritaire tout au long du XIX<sup>e</sup> siècle (Joseph & Newmeyer 2012 : 346). Les extraits suivants sont tirés de Gabelentz 1891 :

*Indeed the Humboldtian view is still haunted by the abrupt dualism between Hellenes and barbarians, people of culture and savages. He speaks of perfect and less perfect languages.* (Gabelentz 1891 : 405, cité et traduit par McElvenny 2017 : 7)

L'auteur tout en prenant en partie ses distances par rapport à la vision humboldtienne des langues postulant une relation indiscutable entre langue et capacités spirituelles, assigne à la linguistique générale la tâche d'étudier les relations entre langue et degré d'évolution de ses locuteurs :

*The goal to which general linguistics must aspire can be no other than to establish the mutual relations between national character and language. Here the mental and temperamental type, the living conditions, the level of civilisation of peoples and families of peoples – there the phenomena, the forces and achievements of their languages. And, between these two, equations that say: the more so on this side, the more or less so on the other.* (Gabelentz 1891: 457, cité et traduit par McElvenny 2017 : 6-7)

*[Gabelentz] carries out several concrete comparisons of structural traits across language groups and links them to the mental and physical conditions of their speakers. The inspiration for this task and the parameters defining it Gabelentz (1891: 372) attributes first and foremost to Humboldt. Given the status Humboldt accords to language as both the product and the continually acting cultivator of the 'national mind' (Volkgeist), a language is to be judged in terms of the way it both reflects and stimulates the thinking of the nation that speaks it.* (McElvenny 2017 : 7)

De telles considérations ont mené un certain nombre de ces philologues et linguistes du XIX<sup>e</sup> siècle à supposer que la complexité morphologique de langues comme le latin, le sanskrit et l'allemand, était l'expression d'une évolution plus avancée, la preuve de civilisations plus prestigieuses.

Le chinois s'est cependant très tôt révélé problématique, comme le montre l'extrait de Renan (1848) sur l'origine du langage, mais avant cela la correspondance entre Humboldt et Abel Rémusat. Celle-ci débute en 1822 (Thouard 1999, 2001)<sup>1</sup> et on y perçoit clairement leur intérêt pour la diversité de ces langues, leur envie de comprendre les différences langagières et de débattre (Thouard 1999 : 14)<sup>2</sup>. Au moment de ce débat, le chinois est l'idiome d'une civilisation encore mystérieuse, une langue opaque, du fait du nombre incroyable de signes qui en composent l'écriture, une énigme :

---

1. Voir aussi Joseph & Newmeyer (2012 : 344-45) sur les échanges entre Humboldt et Rémusat, à propos de la place du chinois dans l'approche pourtant égalitariste de Humboldt : « Toutes les langues peuvent être considérées comme étant au même niveau » (Humboldt cité par Rémusat 1824).

2. Cet échange épistolaire est un modèle de dialogue scientifique comme souligné par Thouard (1999 : 14). Les deux scientifiques dont la formation et l'expérience diffèrent, partagent pourtant un même goût pour la discussion savante, l'argumentation, et savent reconnaître la pertinence de leurs arguments respectifs.

Ce qui amène souvent les linguistes à envisager le monosyllabisme des Chinois comme l'état primitif de toutes les langues, c'est le penchant qui nous porte à regarder la simplicité comme l'indice d'un état d'enfance ou du moins le caractère d'une haute antiquité. Mais c'est là une erreur dont la philologie doit se garder. Le chinois, tout monosyllabique qu'il est, a servi d'organe à une civilisation très développée. [...] Le système grammatical des Hottentots étant beaucoup plus avancé que celui des Chinois, on devrait admettre que les Hottentots ont fait plus de pas que les Chinois dans la voie du développement intellectuel, et sont plus loin de leur état primitif. C'est la une conséquence impossible à soutenir. (Renan 1858 [1848] : 12-13)

La lettre de Humboldt publiée par Rémusat en 1827 fait suite aux échanges entre les deux hommes sur les particularités de cette langue chinoise, mettant à mal les principes de Humboldt :

Cependant le chinois semblait, sous quelques rapports, faire exception au principe de l'auteur [Humboldt], et on appela son attention sur ce singulier phénomène d'un peuple qui, depuis quatre mille ans, possède une littérature florissante, sans formes grammaticales. [...] (Rémusat 1827 : vj)

Difficile en effet de lier simplicité morphologique à faible degré d'évolution culturelle lorsque l'on connaît la grandeur de la civilisation chinoise. La primauté de la grammaire de la tradition occidentale est ici non-pertinente. Même si le chinois n'utilise pas de catégories spécifiques pour marquer les relations entre les mots, les rapports entre les éléments de la langue sont cependant indiqués, comme si la grammaire du chinois était pure syntaxe sans morphologie (Thouard 1999 : 17).

Rémusat dans sa réponse à Humboldt note donc l'incongruité de postuler qu'un « système entier d'interprétations grammaticales [existe] chez un peuple qui n'aurait aucune notion de grammaire » :

[C]ar c'est un fait curieux que la conservation d'un système entier d'interprétations grammaticales chez un peuple qui n'aurait aucune notion de grammaire. (Rémusat 1827 :129)

Ce bref retour dans l'histoire de la typologie morphologique nous a permis de rappeler deux points relatifs à notre propos. Tout d'abord, la complexité a été l'un des critères retenus par les philologues pour un classement des langues et pour statuer sur leur

hiérarchisation. Cependant, les langues isolantes comme le chinois ont mis à mal cette vision du langage et des langues, amenant des philologues comme Gabelentz (1891 : 373) à la fin du XIX<sup>e</sup> siècle à dénoncer une focalisation excessive sur la morphologie ; celle-ci se révélait relativement inopérante pour évaluer le développement d'une langue :

*Gabelentz observes that similar grammatical traits appear to cluster across diverse languages for which none of the usual explanatory factors, such as genealogical relatedness or geographical proximity, can be invoked. [...] Conversely, languages known to be genealogically related frequently differ significantly from one another in their grammatical traits. [...]*

*But, argues Gabelentz (1891: 373), a singleminded focus on morphology as the measure of linguistic development – pursued by Humboldt, Steinthal and August Schleicher (1821–1868) – has proved to be inadequate in the face of the enormous diversity and complexity that has since come to light in the world's languages. (McElvenny 2017 : 14)*

## 2. 2. Complexité et typologie

Avec l'entrée dans le XX<sup>e</sup> siècle, parcouru par une tendance humaniste et égalitaire, les travaux qui associent complexité ou simplicité avec type morphologique de la langue sont rejetés, même si la question reste d'actualité comme en attestent les nombreux travaux contenant le terme « complexité » associé à « langue » (Mufwene, Coupe & Pellegrino 2017). La distinction entre langue complexe et langue simple n'est cependant plus corrélée à une supériorité des unes sur les autres, ou à un niveau de culture. Chez les figures majeures de l'anthropologie linguistique comme Boas ou Sapir, de nombreuses langues primitives sont considérées comme complexes (Boas 1938 [1911] : 160) :

*Both simple and complex types of language of an indefinite number of varieties may be found spoken at any desired level of cultural advance. (Sapir 1921 : 219)*

Ce n'est que dans la deuxième moitié du XX<sup>e</sup> siècle que l'on voit émerger l'idée d'une égale complexité des langues indépendamment de leur type morphologique ou de leur affiliation (Joseph & Newmeyer

2012 : 341)<sup>3</sup>. Cette hypothèse compensatrice, en partie motivée par des considérations idéologiques<sup>4</sup>, restera majoritaire dans le champ de la linguistique jusqu'au début du XXI<sup>e</sup> siècle même si peu de travaux en linguistique définissent le terme ou le concept de complexité de manière explicite. Les premières remises en cause viendront de la sociolinguistique, qui affirme l'importance de facteurs comme le contact de langues, le nombre de locuteurs, le caractère prestigieux ou minorisé d'une langue pour évaluer sa complexité face aux autres langues (Joseph & Newmeyer 2012 : 358) :

*One reason for the dominance of the equi-complexity dogma in the twentieth century, then, was that it meshed well with more modern and egalitarian perspectives, and specifically with the idea that all human speakers are endowed with the same mental, cultural, and biological capacities.* (Szmrecsanyi & Kortmann 2012 : 7)

*Objective measurement is difficult, but impressionistically it would seem that the total grammatical complexity of any language, counting both morphology and syntax, is about the same as any other. This is not surprising, since all languages have equally complex jobs to do, and what is not done morphologically [...] has to be done syntactically.* (Hockett 1958 : 180).

La complexité linguistique s'impose aussi en typologie comme une thématique à explorer. On peut citer ici quelques ouvrages collectifs représentatifs du nouvel attrait pour cette question chez les typologues comme Miestamo, Sinnemäki et Karlsson (2008), Givón & Shibatani (2009), Sampson, Gil et Trugdill, dir. (2009), Ellis & Larsen-Freeman (2009), Coupé, Marsico & Pellegrino (2009) *inter al.*

Cet engouement pour la thématique a été renforcé par la publication de l'article (provocateur) de McWhorter (2001) sur la plus

---

3. Joseph & Newmeyer (2012 : 348) font remonter la première mention explicite de cette égale complexité des langues à 1954.

4. Une motivation de nombreux linguistes, en particulier des descriptivistes, est de montrer que l'humanité n'est pas divisée entre des communautés linguistiques civilisées ayant une pensée complexe, et s'exprimant de façon subtile et des communautés linguistiques primitives incapables de structures langagières élaborées.

grande simplicité de la grammaire des langues créoles<sup>5</sup>. L'intérêt de cet article pour notre propos est qu'il pose la question de la *définition* de la complexité dans les langues mais aussi de son *évaluation*.

Dans cet article, McWhorter postule une plus grande simplicité de la grammaire des langues créoles comme indiqué dans le titre de l'article, en mobilisant l'argument de la profondeur historique qu'il associe avec complexité : la complexité émergerait de façon aléatoire, au fur et à mesure de la lente évolution des langues, pour finalement atteindre un degré de complexité identique dans ces dernières, en vertu des aptitudes et des limites du cerveau humain. Ainsi les créoles, qui sont des langues récentes, n'auraient pas eu le temps d'évoluer vers autant de complexité que des langues plus anciennes. Cette explication cognitive permet à l'auteur de postuler une complexité différenciée selon les langues.

McWhorter s'inscrit aussi en faux contre l'idée d'une égale complexité des langues, qui à son sens, est régulièrement assertée mais n'a jamais été vérifiée de façon systématique et scientifique.

Cependant, rien dans les théories linguistiques modernes ne permet de postuler un *égalisateur* de complexité, un mécanisme qui comparerait les langues, les calibrerait, et les rendrait aussi également complexes. Pourtant, personne ne conteste aujourd'hui que la complexité varie d'un domaine à l'autre selon les langues. Il est facile de montrer la variation des langues en termes de complexité des sous-systèmes, qu'il s'agisse de phonologie ou de morphologie (ex. 1).

Dans le domaine de la phonologie, on pourrait opposer le système consonantique très réduit des langues austronésiennes comme l'hawaïen (8 consonnes) ou des langues papoues comme le rokotas (Maddieson 2011) au système très développé des langues du Caucase ou des langues khoïsanés comme le jullohoansi<sup>6</sup> (89 consonnes) (Dickens 2005).

---

5. D'autres articles publiés dans les années 2000 ont alimenté le débat sur l'égalité de complexité des langues ; on pourrait citer les travaux de Daniel Everett sur la langue pirahã ou ceux de David Gil sur l'indonésien Riau (Voir exemple 14 ci-après).

6. Le jullohoansi est parlé en Namibie et au Botswana (Dickens 2005). Le système phonologique de la langue comprend 89 consonnes dont une cinquantaine de clics.



b.	Mutilek	lagunari	oparia	emango	<b>diote</b>
	garçon.PL.ERG	ami.SG.DAT	cadeau.SG	donner.FUT	AUX.PRES. S3P.O3S.D3S

*Les garçons donneront le cadeau à leur ami*

c.	Mutilak	lagunei	oparia	emango	<b>die</b>
	garçon.SG.ERG	ami.PL.DAT	cadeau.SG	donner.FUT	AUX.PRES. S3S.O3S.D3P

*Le garçon donnera le cadeau à ses amis*

Mais, l'existence de ces variations d'une langue à l'autre ne permet pas de statuer sur une égale complexité des langues en général. McWhorter en vient à conclure que la complexité émerge au fur et à mesure de la lente évolution des langues, les langues atteignant finalement le même degré de complexité. Cette explication cognitive lui permet d'expliquer pourquoi les créoles, qui sont des langues récentes, n'ont pas eu le temps d'évoluer vers autant de complexité que les langues qui existent depuis des siècles :

*We might propose that the volume of such excrescence in each grammar eventually reached the limit of human propensity to process it. [...] Under this scenario all natural languages would be equally complex by virtue of having all come to rest at a certain "surplus complexity quotient".* (McWhorter 2001 :131)

Mais au-delà de cette hypothèse de la compensation, postulée par certains (Riddle 2008 :134-135<sup>8</sup>) et vivement critiquée par d'autres (cf. Shosted 2006<sup>9</sup>, Gil 2008), McWhorter en mettant aussi l'accent

8. Riddle adhère à l'hypothèse de la compensation, selon laquelle si certaines langues (notamment les langues isolantes) sont plus simples morphologiquement, elles sont, par compensation, plus complexes du point de vue syntaxique, sémantique et lexical.

9. Shosted s'appuie sur une analyse statistique du nombre de syllabes et du degré de synthèse des verbes pour démontrer que l'hypothèse d'une complexité compensée au sein d'une langue n'est pas démontrée sur son échantillon de 30 langues : « [...] eleven languages (~34.4%) are [...] both morphologically and phonologically simple (e.g. Koiari, Yoruba, and Amele); seven languages (~21.9%) (e.g. Jullhoansi, Nandi, Vietnamese) are morphologically simple but phonologically complex; seven languages (~21.9%) are both morphologically complex and pho-

sur le consensus non-démonstré, pose indirectement la question de la définition de la complexité et de son évaluation.

### 3. Évaluation de la complexité

Comment définir et évaluer la complexité d'une langue ? Doit-on adopter une vue d'ensemble sur le système langagier ou examiner chaque domaine de la langue séparément ?

Quels critères utiliser pour évaluer cette complexité ?

#### 3. 1. Différentes approches contemporaines de la complexité

Avant de tenter de répondre à la question de l'évaluation de la complexité en linguistique, rappelons les différentes façons de l'appréhender. On peut ainsi évaluer la langue dans son ensemble ou étudier les différents domaines qui la composent séparément. On peut aussi adopter le point de vue de l'apprenant (*user-oriented*), et la complexité sera alors associée au degré de difficulté d'apprentissage. Cette approche « subjective » (*relative approach*) s'oppose à la vision « objective » (*absolute approach*) défendue par les descriptivistes ou grammairiens (Sampson 2009 : 4)<sup>10</sup>, qui associent la complexité d'une langue à la quantité d'informations linguistiques transmises ou à transmettre.

##### 3. 1. 1. Complexité globale vs. complexité spécifique

La complexité globale s'oppose à une complexité spécifique selon que l'on étudie la complexité d'une langue dans son ensemble, ou que l'on évalue les différents domaines de la langue séparément, cette seconde tâche semblant plus réalisable. Szmrecsanyi & Kortmann (2012), Joseph & Newmeyer (2012) citent ainsi de nombreuses études sur la complexité dans des domaines spécifiques. En phonologie, sont

---

*nogically complex (eg. Wichita, Lakhota and Cairene Arabic); and seven (~21.9%) are morphologically complex but phonologically simple (ee.g. Koasati, Yimas, and Ngarinjin).* » (Shosted 2006 : 30)

10. Voir aussi la note de bas de page n°4.

comparés des inventaires de phonèmes ou la structure des groupes consonantiques ; dans le domaine de la morphosyntaxe, c'est le nombre de règles syntaxiques nécessaires à la bonne formation d'une phrase qui est parfois pris en compte<sup>11</sup>.

### 3. 1. 2. Complexité subjective (*user-oriented*) vs. complexité objective (*grammar-oriented*)

La complexité est parfois abordée en termes de degré de difficulté pour les apprenants. Cette approche relative est basée sur l'observation des processus d'apprentissage dans de nombreuses langues. Des études ont en effet montré une similarité dans les processus d'acquisition. Cette approche de la complexité qui évalue la difficulté des locuteurs à encoder ou décoder la langue peut être qualifiée de subjective (*user-oriented*) ; elle s'oppose à une approche plus objective qui évalue la complexité plutôt en termes quantitatifs.

Cette dernière approche (dite aussi *absolute approach*) prend en compte le nombre et la diversité des éléments constituants d'une langue, i.e. les morphèmes (Rescher 2020 [1998] :1), et mesure en termes de longueur les descriptions nécessaires à l'appréhension d'un phénomène (voir Dahl 2004, McWhorter 2001) :

[A relative] approach defines complexity in terms of cost and difficulty to language users, how difficult a phenomenon is to process (encode/ decode) or learn. [...]

The basic idea behind the absolute approach is that the more parts a system has, the more complex it is. To give an example, a language that has 34 phonemes [...] has a more complex phoneme inventory than one that has only 18. (Miestamo 2008 : 24-25).

D'une manière générale et sans entrer dans les détails, l'approche subjective est plus répandue chez les sociolinguistes et psycholinguistes, tandis que l'approche objective est préférée par les typologues.

Cependant, que l'on adopte une vue d'ensemble sur la langue ou qu'on s'intéresse à la complexité d'un domaine spécifique de la

---

11. Joseph & Newmeyer (2012) font ainsi référence aux formules morphologiques proposées par Sapir (1921), aux évaluations métriques des caractéristiques phonologiques de Chomsky & Halle (1968).

langue, que l'on souhaite mesurer des difficultés ou des descriptions de règles morphosyntaxiques, il est impératif de s'appuyer sur des critères d'évaluation valables indépendamment de la langue étudiée, des critères translinguistiques.

### 3. 2. Des critères translinguistiques

The question remains of how degrees of linguistic complexity can actually be measured and compared (Joseph & Newmeyer 2012 : 360)

McWhorter (2005) suggère de mesurer la complexité d'une manière globale. Il propose ainsi un système de mesure basé sur les 3 critères suivants :

- (1) la sur-spécification (*over-specification*),
- (2) le degré d'élaboration des structures langagières (*structural elaboration*),
- (3) le nombre d'irrégularités dans le système (*irregularity*).

En d'autres termes, une langue A sera plus complexe qu'une langue B si elle spécifie des domaines qui sont généralement sous-spécifiés dans les autres langues (i.e. le marquage du duel dans la catégorie « nombre », le marquage de la médiativité ou évidentialité, soit la source de l'information dans le domaine du TAM (Temps-Aspect-Modalité)).

La langue A sera aussi plus complexe que la langue B si le nombre de règles pour passer des formes sous-jacentes aux formes de surface est plus élevé (voir par exemple la morphophonologie des langues inuits). Elle sera aussi plus complexe si elle montre plus d'irrégularités.

Cette proposition de McWhorter a donné lieu à de vifs débats, le système d'évaluation proposé ayant été testé sur d'autres langues que le créole, et certains critères remis en cause. Gil (2008) pointe par exemple la partialité du choix des critères utilisés, pertinents pour démontrer la plus grande simplicité du créole. D'autres auteurs critiquent l'absence de mesures réellement quantifiables, en particulier pour le critère ayant trait aux irrégularités dans le système, qui s'avère être plus qualitatif que quantitatif.

Cet aperçu des différentes approches de la complexité depuis le XIX<sup>e</sup> siècle nous a permis de constater que, qu'il s'agisse de la

classification morphologique des langues chez Humboldt, de la complexité en termes de quantité des descriptions (Dahl 2004) ou du système de mesure proposé par McWhorter (2005), la complexité grammaticale d'une langue y est toujours mesurée en termes de distinctions grammaticales exprimées, ce qui pour le domaine de la morphologie va souvent se traduire par le fait d'associer complexité avec flexion et caractère obligatoire d'une expression. Or les langues d'Asie du Sud-Est et d'Asie orientale, à la morphologie pauvre, nous racontent une autre histoire. Elles suggèrent que cette « complexité exprimée » (*overt complexity*) n'est qu'un type particulier de complexité, l'examen de ces langues asiatiques donnant à voir une complexité invisible, cachée (*hidden complexity*) généralement ignorée dans la littérature.

#### 4. Complexité sous-jacente, une particularité structurelle des langues d'Asie du Sud-est et d'Asie orientale

Cette complexité cachée prend racine dans une particularité des langues d'Asie, qualifiée d'indétermination (*indeterminateness*) par Bisang (2004, 2009). Cette indétermination se manifeste par (a) l'absence d'expressions linguistiques correspondant aux arguments, et par (b) l'absence ou la rareté des catégories grammaticales obligatoires. Ces deux caractéristiques entraînent une importance accrue du contexte pour l'interprétation d'un énoncé, un fait déjà noté par les missionnaires du XVIII<sup>e</sup> siècle confrontés au chinois, comme le montre cet extrait d'un père missionnaire datée de 1769 cité par Thouard 1999 :

« [...] le chinois est bien difficile. [...] Le même mot n'a jamais qu'une terminaison ; on n'y trouve point du tout ce qui dans nos déclinaisons distingue le genre et le nombre des choses dont on parle. Dans les verbes, rien ne nous aide à entendre quelle est la personne qui agit, comment et en quel temps elle agit, si elle agit seule ou avec d'autres. En un mot, chez les chinois le même mot est substantif, adjectif, verbe, adverbe, singulier, pluriel, masculin, féminin, etc. C'est à vous qui écoutez, à épier les circonstances et à deviner. » [...]

L'impression laissée par cette description est que le chinois ne peut s'apprendre qu'in situ. [...] Les pères sont en droit de se plaindre : il leur faut apprendre une langue dans laquelle les mots sont flottants, n'ont pas de détermination fixe. C'est toujours à l'auditeur ou au lecteur à compléter les relations à actualiser la grammaire. (Thouard 1999 : 12)

Une troisième manifestation de cette indétermination est la grande polyfonctionnalité des formes. Celles-ci ne sont pas assignées de façon rigide à une catégorie et peuvent apparaître dans des contextes variés ; leur fonction dépend alors du contexte grammatical :

*The morphemes used to denote grammatical functions are not limited to one particular category exclusively. Depending on the context, they can express a variety of grammatical categories without covering the whole functional range of each of these categories. My hypothesis is that this is an areal phenomenon of East and mainland Southeast Asian languages (Bisang 2004 : 109-10).*

Ces différentes caractéristiques des langues d'Asie (Vittrant 2010) seront illustrées dans les sections suivantes.

#### 4. 1. Présence des arguments non-obligatoire (*pro-drop* à l'extrême)

Les langues d'Asie du Sud-Est manifestent une tendance marquée à l'ellipse des constituants et une absence d'indexation grammaticale de ces derniers dans la sphère verbale. Cette absence des arguments est structurelle. Elle implique une reconstruction des informations manquantes par la situation d'interlocution.

L'exemple birman (3) montre une absence totale de formes faisant référence aux participants à l'action, qu'il s'agisse de pronoms ou d'indice personnel sur les verbes. Cette absence ne doit pas laisser penser que les formes personnelles n'existent pas en birman. Au contraire, les pronoms sont nombreux dans cette langue (Vittrant 2019 : 84) quoique non-obligatoires, à la différence des pronoms français ou des marques de personnes de langues comme l'espagnol.

Birman (Tibéto-birman, Birmanie/Myanmar)

(3) ပြင်ပေးလိုက်မယ်။

pyiN <sup>2</sup>	pe <sup>3</sup>	lai <sup>?</sup>	=me <sup>2</sup>
réparer	donner/BENEF	suivre/ASP:TERM	IRR
[ʃe]	[le]	réparerai pour [vous].	

Comme le montre cet exemple, un énoncé bien formé en birman peut être composé d'un complexe verbal seul sans aucune expression linguistique faisant référence aux participants de l'action.

#### 4. 2. Absence ou rareté des catégories grammaticales obligatoires

Une seconde caractéristique des langues d'Asie du Sud-Est et sinitiques est la tendance de ces langues à être isolantes<sup>12</sup>. Cette grande simplicité morphologique se traduit par une absence de flexion grammaticale pour le nombre, la personne, et une absence de marqueurs temporels obligatoires.

L'exemple (4a) illustre l'absence de *catégorie du nombre* obligatoire en birman : les trois syntagmes nominaux sont non-marqués pour le nombre malgré l'interprétation plurielle attendue dans le contexte pour les deux premiers. De même en (b), la pluralité sera inférée du contexte en l'absence de marque dédiée. Pourtant ces marques existent ; elles ne sont cependant pas obligatoires dans les cas d'arguments non spécifiques (Vittrant, 2022).

Birman (Tibéto-birman, Birmanie/Myanmar) – d'après Bernot *et al.* (2001: 93)

(4) a. စားပွဲကုလားထိုင်အစားဖျာခင်းမယ်။

zəbwε <sup>3</sup>	kəla <sup>3</sup> .thaiN <sup>2</sup>	ʔə.sa <sup>3</sup>	phyə <sup>2</sup>	khiN <sup>3</sup>	=Me <sup>2</sup>
table	indien.siège	à.la.place.de	natte	étaler	=IRR
(On) étalera <u>une</u> natte à la <u>place</u> <u>des</u> tables et <u>des</u> chaises.					

12. Pour certaines familles de langues comme la famille sino-tibétaine, il est proposé de reconstruire un certain degré de morphologie, sur la base des langues actuelles morphologiquement très complexes comme les langues rgyalroniques. Voir DeLancey (1989, 2010), Jacques (2016) d'une part et LaPolla (1992) d'autre part, pour des discussions sur la complexité morphologique du proto-sino-tibétain.

b. သူ ရန်ကုန်မှာ ဝယ်မှာက စာအုပ်နဲ့ခဲတံပါ။

θu<sup>2</sup> Yan<sup>2</sup>gon<sup>2</sup> =Ma<sup>2</sup> wε<sup>2</sup> Ma<sup>2</sup> =Ka<sup>1</sup> sa<sup>2</sup>ʔoʔ =nε<sup>1</sup> khε<sup>3</sup>daN<sup>2</sup> =Pa<sup>2</sup>  
 3SG Yangon LOC acheter NMLZ.IRR =TOP livre(s) =avec/et crayon(s) =POL  
*Ce qu'il va acheter à Rangoun, (ce sont) (des) livres et (des) crayons.*

Les exemples (5) et (6) illustrent le caractère invariable des verbes non marqués pour le temps. (Notez aussi l'absence d'information concernant le nombre). Comme l'indiquent les différentes traductions proposées pour l'énoncé (5a), le thaï ne nécessite pas de spécifier grammaticalement le temps : les interprétations passée ou présente sont toutes deux possibles et acceptables. Il est néanmoins possible de désambiguïser cet énoncé, soit en introduisant un complément circonstanciel de temps comme en (5b), soit en spécifiant le caractère accompli de l'action (5c), ce qui force alors la lecture « passé » de l'événement.

Thaï (Tai-kadai, Thaïlande) d'après Bunkham (communication personnelle)

- (5) a. khǎw **àan** nǎngsǐi b. toonnú khǎw **àan** nǎngsǐi  
 3SG **lire** livre maintenant 3SG **lire** livre  
 i. *Il/ elle lit un livre* *Il/ elle lit maintenant/en ce moment*  
 ii. *Il/ elle a lu un livre*
- c. khǎw **àan** nǎngsǐi léew  
 3SG **lire** livre ASP : CRS  
*Il/ elle a lu un livre*

Comme précédemment pour le thaï, l'indétermination temporelle est levée, en l'absence de marque grammaticale, par la présence de circonstanciels en birman (6).

Birman (Tibéto-birman, Birmanie/Myanmar)

- (6) a. ဧပြီလမှာနေ့ပူတယ်။  
 ʔε<sup>2</sup>pyi<sup>2</sup>.la<sup>1</sup>=ma<sup>2</sup> **ne<sup>2</sup> - pu<sup>2</sup>** =Tε<sup>2</sup>  
 Avril.mois=LOC **soleil- être.chaud** =REAL  
*En avril, il fait chaud. [présent générique]*

- b. လွန်ခဲ့တဲ့လကလဲနေပူတယ်။  
 lwaN<sup>2</sup>.Khe<sup>1</sup>.Te<sup>2</sup>.la<sup>1</sup>=Ka<sup>1</sup> le<sup>3</sup>      **ne<sup>2</sup> - pu<sup>2</sup>**      =Te<sup>2</sup>  
 dépassé.mois      =TOP aussi      **soleil- être.chaud**      =REAL  
*Le mois dernier aussi, il faisait chaud.* [passé]

Le syntagme verbal est identique dans les deux énoncés, et ne comprend pas de marques temporelles<sup>13</sup>. Seule la présence d'un complément circonstanciel temporel, i.e. « en avril » pour l'énoncé (a) et « le mois dernier » pour l'énoncé (b), permet d'ancrer temporellement le procès.

#### 4. 3. Importance de la pragmatique et du contexte

Un effet de cette indétermination structurelle est l'importance du composant pragmatique ; en effet, la structure de ces langues ne force pas le locuteur à exprimer certaines notions à l'aide de matériel linguistique (affixes, morphèmes, mots, syntagmes...) lorsque ces informations peuvent être inférées du contexte. Cette absence d'information grammaticale obligatoire a pour conséquence une importance accrue du contexte pour le décodage d'un énoncé. En d'autres termes, la présence d'un contexte discursif est primordiale pour analyser correctement un énoncé dans ces langues

Les exemples (7) et (8) illustrent le caractère éminemment pragmatique d'une langue comme le birman ou le stieng. Dans ces langues, un référent introduit à l'aide de matériel linguistique en début de discours n'apparaît plus ni sous une forme lexicale ni par une reprise anaphorique tant qu'il reste actif et accessible à l'interlocuteur.

En (7) est présenté un extrait de monologue enregistré dans la région du Mont Popa en Birmanie, qui traite des spécificités des *nats* – des esprits tutélaires – habitant la région. Les phrases en (a) et (b) se suivent dans le discours.

13. Pour plus de précision sur la grammaire des verbes birmans, voir Vittrant 2004, 2012, 2019.

## Birman (Tibéto-birman, Birmanie/Myanmar)

(7) a. နတ်ကမကြိုက်တာ ၊ အနက် ၊ ဘာကြောင်းလဲဆိုတော့ ၊  
 naʔ =Ka<sup>1</sup> mə=Caiʔ =Ta<sup>2</sup> / əʔ.neʔ / ba<sup>2</sup>.CoN<sup>3</sup> le<sup>3</sup> sho<sup>2</sup> =Tə<sup>1</sup>  
 nat (esprit) s./TOP<sup>14</sup> NEG aimer REAL noir (N.) pourquoi QST dire TOP  
 အဲဒါတော့ကျွန်တော်လဲမသိဘူး။  
 ʔε<sup>3</sup>.da<sup>2</sup>=Tə<sup>1</sup> cəno<sup>2</sup> le<sup>3</sup> mə= θi<sup>1</sup> =Phu<sup>3</sup>  
 DEM.ANAPH=TOP 1SG(H) aussi NEG savoir NEG  
*Les nats, pourquoi (ils) ne (l') aiment pas, la (couleur) noire ? Si on se  
 demande pourquoi, et bien ça, je ne le sais pas !*

b. ဒါပေမဲ့မကြိုက်တာတွေမှန်တယ်။  
 da<sup>2</sup>pe<sup>2</sup>me<sup>3</sup> Ø Ø mə= Caiʔ =Ta<sup>2</sup> =Tə<sup>1</sup> məN =Te<sup>2</sup>  
 mais [ils] [noir] NEG aimer REAL.NF TOP ê. vrai REAL  
*Mais qu'(ils) n'aiment pas (le noir), ça c'est vrai !*

En (7a), l'expression renvoyant aux esprits ('nats') apparaît marquée par une particule (က /Ka<sup>1</sup>/) qui indique que le référent est le topique et à la source de l'action d'« aimer ». L'expression renvoyant à ce qui n'est pas aimé fait l'objet d'un détachement à droite, et apparaît sous la forme du lexème nominal signifiant « le noir ».

La phrase (b), qui suit dans le monologue, ne contient ni pronom anaphorique, ni expression lexicale renvoyant aux protagonistes du procès d'aimer, ni indice personnel sur le verbe. En d'autres termes, les informations sont linguistiquement manquantes et doivent être inférées par l'interlocuteur à partir du contexte d'interlocution.

La même stratégie est observable en stieng, comme illustré par l'exemple (8) suivant.

## Stieng (Mon-khmer, Cambodge) – d'après Bon (2014)

- (8) a. koəndren sə:h sala: m=bu: ; paŋ kan toəŋ kələʔ bə:h pəjkruc  
 enfant écolier un=CLF.PERSONNE 3SG tenir perche cueillir orange  
*L'écolière (seule) tient une perche pour cueillir une orange.*

14. Le morphème grammatical =Ka<sup>1</sup> est polyfonctionnel. Il pourra marquer la source (ablatif, nominatif) d'un procès, ou l'argument topique, en particulier dans les énoncés monoactanciels qui ne nécessitent pas de marquage syntaxique (rôle des arguments non-ambigus). Il sera donc glossé <S.> ou <top> et parfois les deux quand il n'est pas possible de trancher hors contexte.

b. təəm kruc      niə                      pɛj      ʔək                      mət  
 oranger          DEM.DIST              fruit      ê.nombreux          très  
*Cet oranger [a] de très nombreux fruits.*

c. bɛ:h    ba:n      hə:j,    bə:h-tət-ʔəw,    paŋ    pək      hə:j,    paŋ    sa: [...]   
 cueillir RESULT PERF 'à présent' 3    éplucher PERF 3    manger  
*Une fois qu'[elle] a cueilli [l'orange], elle [l']épluche, et puis elle [la] mange ; [...]*

Les différents participants aux actions décrites dans la phrase (c), i.e. l'écolière et l'orange, sont précédemment introduits par des expressions lexicales (phrase (a)). Éléments du discours toujours présents et accessibles à la conscience de l'interlocuteur, ils peuvent facilement être inférés du contexte et ne nécessitent pas d'être rappelés par une forme linguistique dans les propositions suivantes.

#### 4. 4. Polyfonctionnalité des formes

Une dernière manifestation de cette indétermination des langues d'Asie est la polyfonctionnalité des formes. Il est aisé de montrer leur hétérosémie (Lichtenberk 1991)<sup>15</sup> ; ces formes, utilisées dans différents environnements, peuvent remplir différentes fonctions —comme illustré par le vietnamien en (9). Elles peuvent aussi être associées à différentes valeurs sémantiques, rendues dans nos langues par des marqueurs distincts comme dans les exemples birmans (10) & (11).

Dans l'exemple (9) du vietnamien, le morphème *thành công* véhiculant la notion de 'réussite' s'interprétera alternativement comme un nom, un verbe ou un adjectif selon le contexte, sans que la forme ne subisse de modification.

---

15. L'hétérosémie est définie comme un cas particulier de polysémie, pour lequel les différents sens apparentés d'un morphème sont associés à des contextes grammaticaux distincts. En d'autres termes, le sens d'une forme est inféré de son environnement syntaxique. Cette notion n'est pas limitée à la grammaticalisation ; elle peut aussi décrire les alternances de formes verbales et nominales similaires, et s'inscrit dans une approche syntactico-pragmatique. Voir Lichtenberk (1991) pour le concept et Enfield (2006) pour son application aux langues d'Asie du Sud-Est.

Vietnamien (Mon-khmer, Vietnam) – d’après Do-Hurinville (2010)

(9) a. Paul đã **thành.công**  
 Paul ACCOMP réussir  
*Paul a réussi.*

b. Bài diễn.văn của Paul rất **thành.công**  
 CLF discours POSS Paul très être réussi  
*Le discours de Paul est très réussi.*

c. **Thành.công** của dự.án...  
 réussite POSS projet  
*La réussite du projet...*

Parfois la prosodie, composante essentielle du contexte discursif, peut se révéler déterminante pour l’interprétation comme en (10). Dans cet exemple, les syllabes accentuées sont indiquées en petites capitales et en gras. Le morphème <con>, dont le sens lexical est « enfant », est associé à différentes catégories de mots selon qu’il est ou non accentué.

Vietnamien (Mon-khmer, Vietnam) – d’après Do-Hurinville (2008)

(10) a. **XE** con hỏng rồi.  
**voiture 1SG** être en panne ASP  
*Ma voiture est tombée en panne.*

b. **XE CON** hỏng rồi.  
**voiture enfant** être en panne ASP  
*La voiture de notre enfant est tombée en panne.*

c. Xe **CON** hỏng rồi.  
 voiture **ê. petite** être en panne ASP  
*La petite voiture est tombée en panne.*

Postposé au nom et non-accentué en (a), *con* est interprété comme un pronom de 1<sup>ère</sup> personne, la marque d’un locuteur socialement inférieur à son interlocuteur. En (b), il a son sens lexical d’origine (« enfant ») quand le syntagme nominal dans lequel il apparaît est accentué. En (c), *con* est le seul morphème accentué ; il qualifie le nom

qu'il suit, et sera interprété comme ayant la propriété d'être « comme un enfant », i.e. « petit ». La forme *con* est ainsi analysée comme relevant de différentes parties du discours selon l'environnement syntaxique et la prosodie.

De même, le morphème ရ /ya<sup>1</sup>/ signifiant « *obtenir, acquérir* » en birman, et illustré par l'exemple (11), pourra recevoir d'autres interprétations en fonction du contexte discursif et de l'environnement syntaxique.

Birman (Tibéto-birman, Birmanie/Myanmar)

(11) သစ်ခိုတ်သမားသာဆုငွေရပါတယ်။  
 θəkhoʔ - θəma<sup>3</sup> θa<sup>2</sup> shu<sup>1</sup>-ŋwe<sup>2</sup> **ya<sup>1</sup>** =Pa<sup>2</sup> =Tε<sup>2</sup>  
 bûcheron seulement récompense **obtenir** POL REAL  
*Seul le bûcheron a touché (/ touche) la récompense.*

L'extrait présenté en (12) provient d'un texte décrivant la vie en Birmanie au début du xx<sup>e</sup> siècle. Les deux phrases se suivent dans le texte original. Elles contiennent toutes deux le même verbe သုံး /θoN<sup>3</sup>/ « utiliser » suivi du verbe ရ /ya<sup>1</sup>/ « obtenir ». Comme souvent dans les constructions verbales en série, l'un des verbes peut évoluer pour devenir marqueur grammatical (Vittrant 2006). C'est le cas du morphème signifiant « acquérir », qui en birman comme dans de nombreuses langues d'Asie, prend une valeur modale (Van der Auwera *et al.* 2009). Notez cependant que les traductions de (a) et (b) donnent des interprétations différentes de cette séquence avec marqueur modal, distinguant entre deux notions sémantiques importantes à nos yeux, à savoir la possibilité et la nécessité.

(12) (a) ကျုံးပတ်လည်းကအရပ်တွေကကျုံးရေကိုသုံးရတယ်။ (Vittrant 311 : 2004)

coN <sup>3</sup>	paʔ-le <sup>2</sup>	=Ka <sup>1</sup>	ʔə-yaʔ-Twe <sup>2</sup>	=Ka <sup>1</sup>
douve	entourer- tourner (n.)	S.	quartier - PL	S.
coN <sup>3</sup>	ye <sup>2</sup>	=Ko <sup>2</sup>	θoN <sup>3</sup>	<b>ya<sup>1</sup></b> =Tε <sup>2</sup>
douve	eau	OBJ	utiliser	<b>MOD</b> REAL

[choix] *Les quartiers entourés de douves avaient la possibilité d'utiliser l'eau des douves.*

- (b) ကျန်ရပ်ကွက်တွေအားလုံးက တွင်းရေကို ခပ်သုံးရတာ ၊
- |  |                                  |                  |                                    |                  |
|--|----------------------------------|------------------|------------------------------------|------------------|
| caN <sup>2</sup> -yaʔkweʔ-Twe <sup>2</sup> | ʔa <sup>3</sup> loN <sup>3</sup> | =Ka <sup>1</sup> | twiN <sup>3</sup> -ye <sup>2</sup> | =Ko <sup>2</sup> |
| ê. de reste-quartier-PL                    | tous                             | s                | puits-eau                          | OBJ              |
| khaʔ                                       | θoN <sup>3</sup>                 | ya <sup>1</sup>  | =Ta <sup>2</sup>                   |                  |
| puiser                                     | utiliser                         | <b>MOD</b>       | REAL.NF                            |                  |

[contrainte] *Tous les quartiers restants devaient puiser et utiliser l'eau des puits.*

Ainsi en (a), les habitants des quartiers des douves sont face à un choix, alors qu'en (b) les autres habitants subissent une contrainte, celle d'aller au puits... ce que nous traduisons en français par deux morphèmes modaux distincts, i.e. « pouvoir » vs. « devoir ».

La fonction modale du morphème ရှိ /ya<sup>1</sup>/ apparaît encore dans l'exemple (13), dont l'interprétation dépendra du contexte.

- (13) ဘယ်လမ်းကသွားရမလဲ။ (Vittrant, 312 : 2004)

be <sup>3</sup>	laN <sup>3</sup>	=Ka <sup>1</sup>	θwa <sup>3</sup>	ya <sup>1</sup>	=mə	=lɛ <sup>3</sup>
PR.QST	route	s.	aller	MOD	IRR.QST	QST

(a) *Quel chemin doit-on prendre ? (nécessité)*

(b) *Quel chemin peut-on prendre ? (possibilité)*

L'exemple (14) ci-dessous résume bien la problématique de l'indétermination des langues d'Asie du Sud-Est en montrant comment l'absence de catégories requises peut engendrer des interprétations variées ; la prise en compte du contexte est alors nécessaire pour obtenir la complétude sémantique attendue.

Cet énoncé d'indonésien est constitué de deux mots référentiels, *ayam* et *makan* ; il ne contient aucun marqueur grammatical. Le caractère vague – plutôt qu'ambigu – de la séquence *ayam makan*, est rendu par une traduction très générale qui permet toutes les interprétations. Cinq traductions plus précises sont cependant proposées par Gil (2008), reflétant des constructions sous-jacentes distinctes.

Le caractère mono-morphémique de chacun des mots illustre aussi la morphologie isolante de la langue, tandis que leur comportement grammatical identique (catégorie grammaticale non-spécifiée) produit les interprétations multiples de la séquence.

Indonésien Riau (austronésienne, Indonésie) – d'après Gil (2008 : 114)

(14) ayam            makan  
       poulet        mange(r)

*'quelque chose qui a à voir avec poulet et manger'*

- a. (X) mange(nt) du poulet
- b. Le(s) poulet(s) mange(nt) / ont mangé
- c. Un/Des poulet mange(nt) / ont mangé
- d. Le poulet est mangé / a été mangé
- e. Des poulets sont mangés

Pour résumer, l'indétermination est une caractéristique des langues d'Asie – en particulier d'Asie du Sud-Est et des langues sinitiques – qui se manifeste par une certaine polyfonctionnalité ou hétérosémie. Une même forme peut ainsi encoder différentes significations lexicales et grammaticales, appartenir à différentes parties de discours, et participer à des constructions distinctes. Le caractère vague d'un énoncé sera alors compensé par des informations reconstruites du contexte discursif (i.e. prosodie) et de l'environnement syntaxique (i.e. constructions sous-jacentes).

## 5. Conclusion

Dans cet article, après un rapide historique de la notion de complexité en linguistique, nous avons souhaité interroger les méthodes d'évaluation du phénomène, l'hypothèse de la compensation étant trop souvent posée comme un consensus non-démonstré dans la littérature.

Différentes approches ont été présentées, abordant la complexité du point de vue de l'apprenant, ou d'un point de vue plus objectif à partir de données quantitatives (quantité d'éléments dans un domaine, quantité d'information à transmettre...). Nous avons aussi examiné les critères d'évaluation de la complexité d'une langue proposés par McWhorter (2005). Tous ces travaux ont un point

commun : ils ne tiennent compte que de la complexité exprimée et visible et continuent d'associer un fort degré de complexité à une morphologie flexionnelle (i.e. nombre et variété des formes grammaticales obligatoires). Les langues d'Asie d'aujourd'hui à tendance isolante comme le chinois mandarin ou le thaï, sont alors considérées comme peu complexes.

Or le caractère éminemment *pro-drop* de ces langues (absence d'indices personnels), l'absence d'ancrage temporel obligatoire, leur fonctionnement éminemment pragmatique (polyfonctionnalité, hétérosémie, interprétation basée sur le contexte) reflètent une complexité d'un autre genre, une complexité « cachée » largement illustrée dans cet article. Ces particularités linguistiques posent la question de la pertinence des critères utilisés précédemment pour mesurer la complexité dans une langue.

En d'autres termes, la simplicité des formes de surface dans beaucoup de langues asiatiques n'est pas synonyme d'absence de complexité comme on pourrait le supposer en se basant sur la définition de la complexité proposée par Rescher (2020 [1998]) ou sur des critères du type de ceux qui ont été proposés par McWhorter (2005). Il est donc essentiel de retravailler en premier lieu à la définition des critères permettant de mesurer la complexité d'une langue.

## Bibliographie

- Bernot Denise, Marie-Hélène Cardinaud & Marie Yin Yin Myint. 2001. *Grammaire Birmane - Manuel du Birman*. vol. 2. Paris : L'Asiathèque (Langues et Mondes).
- Bisang, Walter. 2004. Grammaticalization without coevolution of form and meaning: The case of tense-aspect-modality in East and mainland Southeast Asia. *What Makes Grammaticalization? - A Look from its Fringes and its Components*, dir. par Walter Bisang, Nikolaus Himmelmann, N. & Björn Wiemer. Berlin : Mouton de Gruyter. 109- 138.
- Bisang, Walter. 2009. On the evolution of complexity—sometimes less is more in East and mainland Southeast Asia. *Language Complexity as an Evolving Variable*, dir. par Geoffrey Sampson, David Gil & Peter Trudgill. Oxford: Oxford University Press. 34-49

- Boas, Franz. 1938 [1911]. *The Mind of Primitive Man*. New York: Macmillan.
- Bon, Noëllie. 2014. *Une grammaire de la langue stieng, langue en danger du Cambodge et du Vietnam*. Thèse de Doctorat. Lyon : Université Lumière Lyon 2.
- Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern in English*. New York : Harper & Row
- Coupé, Christophe, Egidio Marsico & François Pellegrino. 2009. Structural complexity of phonological systems. *Approaches to Phonological Complexity*, dir. par Pellegrino, F., Marsico, E., Chitoran, I. & Coupé, C., Berlin, New York: Mouton de Gruyter (Phonology & Phonetics Series, 16). 141-169
- Creissels, Denis. 2006. *Syntaxe générale, une introduction typologique 1*. Paris : Lavoisier.
- Dahl, Östen. 2004. *The Growth and Maintenance of Linguistic Complexity*. Amsterdam-Philadelphie : Benjamins.
- DeLancey, Scott. 1989. Verb agreement in Proto-Tibeto-Burman. *Bulletin of the School of Oriental and African Studies, University of London*, 52(2). 315-333.
- DeLancey, Scott. 2010. Towards a history of verb agreement in Tibeto-Burman. *Himalayan Linguistics* 9(1). 1-39.
- Dickens, Patrick J. 2005. *A Concise Grammar of Ju|'hoan, with a Ju|'hoan-English Glossary and a Subject Index*. Cologne : Rüdiger Köppe (Quellen zur Khoisan-Forschung, 17).
- Do-Hurinville, Thành. 2006. Étude de quelques coverbes, de l'ordre temporel et du discours rapporté dans la littérature et dans la presse vietnamiennes, Étude contrastive avec le français. *Bulletin de la Société de linguistique de Paris* 101(1). 369-416.
- Do-Hurinville, Thành. 2008. Quelques remarques sur les relations syntaxiques, sémantiques et accentuelles en vietnamien. *Travaux linguistiques du CerLiCo 21 [Grammaire et prosodie]*. 183-196.
- Do-Hurinville, Thành. 2010. Les parties du discours en vietnamien, grammaticalisation et transcatégorialité. *Bulletin de la Société de linguistique de Paris* 105(1). 327-370
- Ellis, Nick C., & Larsen-Freeman, D., éd. 2009. *Language as a Complex Adaptive System*. Mahwah, NJ : Wiley.
- Enfield, Nick. J. 2006. Heterosemy and the grammar-lexicon trade-off. *Catching Language*, dir. par Felix K. Ameka, Alan Dench, & Nicholas Evans. Berlin : Mouton de Gruyter. 297-320.
- Givón, Talmy & Shibatani, Masayochi. 2009. *Syntactic Complexity: Diachrony, Acquisition, Neuro-cognition, Evolution*. Amsterdam : Benjamins.

- Gabelentz, Georg von der. 1891. *Die Sprachwissenschaft: ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Leipzig : T.O. Wigiel Nachfolger
- Gil, David. 2008. How complex are isolating languages? *Language Complexity: Typology, Contact, Change*, dir. par Matti Miestamo, Kaius Sinnemäki & Fred Karlsson. Amsterdam, Philadelphie : John Benjamins. 109-131.
- Hovelacque, Abel & Julien Vinson. 1878. *Études de linguistique et d'ethnographie*. Paris : C. Reinwald & Cie Libraires-editeurs.
- Hockett, Charles F. 1958. *A course in Modern Linguistics*. New York : Macmillan.
- Jacques, Guillaume. 2016. Le sino-tibétain : polysynthétique ou isolant ? *Faits de langues* 47(1). 61-74
- Joseph, John & Frederick Newmeyer. 2012. All languages are equally complex. The rise and fall of a consensus. *Historiographia Linguistica* 39(2-3). 341-368
- LaPolla, Randy. 1992. On the dating and nature of the verb agreement in Tibeto-Burman. *Bulletin of the School of Oriental and African Studies* 55(2). 298-315.
- Lichtenberk, Frantisek. 1991. Semantic Change and Heterosemy in Grammaticalization. *Language* 67(3). 475-509.
- Maddieson, Ian. 2011. Consonant Inventories. *The World Atlas of Language Structures Online*, dir. par Matthew S. Dryer & Martin Haspelmath. Munich : Max Planck Digital Library. [<http://wals.info/chapter/1>, consulté le 16/02/2023].
- McElvenny, James. 2017. Grammar, typology and the Humboldtian tradition in the work of Georg von der Gabelentz. *Language & History* 60(1). 1-20.
- McWhorter, John H. 2001. The world's simplest grammars are creole grammars. *Linguistic Typology* 5. 125-166.
- McWhorter, John H. 2005. *Defining Creole*. Oxford : Oxford University Press.
- Miestamo, Matti. 2008. Grammatical complexity in a cross-linguistic perspective. *Language Complexity: Typology, Contact, Change*, dir. par Matti Miestamo, Kaius Sinnemäki & Fred Karlsson. Amsterdam: Benjamins. 23-41.
- Miestamo, Matti, Kaius Sinnemäki & Fred Karlsson, dir. 2008. *Language Complexity: Typology, Contact, Change*. Amsterdam : Benjamins.
- Mufwene, Saliko, Christophe Coupé & François Pellegrino. 2017. Complexity in Language: A Multifaceted Phenomenon. *Complexity in Language: Developmental and Evolutionary Perspectives*, dir. par Saliko Mufwene, Christophe Coupé & François Pellegrino. Cambridge : Cambridge University Press. 1-29.

- Renan, Ernest. 1858 [1848]. *De l'origine du langage*. Paris : Michel Levy Frères.
- Rescher, Nicholas. 2020 [1998]. *Complexity : a philosophical overview*. New York : Routledge
- Riddle, Elizabeth M. 2008. Complexity in isolating languages: Lexical elaboration versus grammatical economy. *Language Complexity: Typology, Contact, Change*, dir. par Matti Miestamo, Kaius Sinnemäki, & Fred Karlsson, Amsterdam, Philadelphie : Benjamins. 133-151
- Rémusat, Jean-Pierre Abel. 1824. Compte-rendu de 'Sur la naissance des formes grammaticales' de Humboldt (1821). *Journal Asiatique* 5. 51-61.
- Rémusat, Jean-Pierre Abel. 1827. *Lettre à Rémusat, sur la nature des formes grammaticales en général, et sur le génie de la langue chinoise en particulier*. Paris : Librairie Orientale de Dondey-Dupré père et fils.
- Sampson, Geoffrey. 2009. A linguistic Axiom Challenged. *Language Complexity as an Evolving Variable*, dir. par Geoffrey Sampson, David Gil & Peter Trudgill. Oxford, New York : Oxford University Press. 1-18.
- Sampson, Geoffrey, David Gil & Peter Trudgill, dir. 2009. *Language Complexity as an Evolving Variable*. Oxford, New York : Oxford University Press.
- Sapir, Edward. 1921. *Language: An introduction to the Study of Speech*. New York : Harcourt & Brace
- Shosted, Ryan. 2006. Correlating complexity: A typological approach. *Journal of Linguistic Typology* 10. 1-40.
- Szmrecsanyi Benedikt & Bernd Kortmann. 2012. *Linguistic Complexity: Second Language Acquisition, Indigenization, Contact*. Berlin : De Gruyter,
- Thouard, Denis. 1999. Humboldt, Abel-Remusat et le chinois : la recherche de la correspondance. *Lettres édifiantes et curieuses sur la langue chinoise*, dir. par Jean Rousseau & Denis Thouard. Lille : Septentrion. 9-28.
- Thouard, Denis. 2001. Humboldt, Abel-Rémusat et le chinois. Du mystère au savoir. *Texto !* [[http://www.revue-texto.net/Inedits/Thouard\\_Humboldt.html](http://www.revue-texto.net/Inedits/Thouard_Humboldt.html)], consulté le 16/02/2023].
- Van der Auwera van der Auwera, Johan, Peter Kehayov & Alice Vittrant. 2009. Acquisitive modals. *Cross-linguistic Studies of Tense, Aspect, and Modality*, dir. par Lotte Hogeweg, Helen de Hoop & Andrej Malchukov. Amsterdam : Benjamins. 271-302.
- Vittrant, Alice. 2004. *La modalité et ses corrélats en birman dans une perspective comparative*. Thèse de Doctorat. Université Paris 8 – Vincennes/Saint-Denis.
- Vittrant, Alice. 2006. Les constructions en série : une autre approche du syntagme verbal en birman. *Bulletin de la Société de linguistique de Paris* 101(1). 305-368.

- Vittrant, Alice. 2010. Aire linguistique Asie du Sud-Est continentale : le birman en fait-il partie ? *Moussons* 16. 7-38.
- Vittrant, Alice. 2012. Définir la modalité : vers une théorie linguistique de la modalité à partir de son expression dans les langues. *Recueil en hommage à Robert Vion*, dir. par Claire Maury-Rouan. Aix-en-Provence : Presses Universitaire de Provence. 107-133.
- Vittrant, Alice. 2019. Burmese. *The Mainland Southeast Asian Linguistic Area*, dir. par Alice Vittrant & Justin Watkins. Amsterdam : Mouton de Gruyter. 56-130.
- Vittrant, Alice. 2022. Où est l'accord dans une langue sans accord ? L'exemple du birman. *Travaux linguistiques du CERLICO - Accord non-accord*. Rennes : Presses Universitaires de Rennes. 175-195.
- Vittrant, Alice & Justin Watkins, dir. 2019. *The Mainland Southeast Asian Linguistic Area*. Amsterdam : Mouton de Gruyter.

# Le défaut de « forme interne » dans les langues mandingues selon la psychologie des peuples de Steintal (1867)

**Jacques François**

Université de Caen-Normandie – CRISCO, EA 4255

---

## RÉSUMÉ

En 1848, Steintal a esquissé une vaste entreprise épistémologique fusionnant la vision de Humboldt sur le génie des langues (*Sprachgeist*) indissociable de l'esprit des peuples qui les parlent (*Volksgeist*) et celle de Hegel sur la phénoménologie de l'esprit progressant dialectiquement d'âge en âge. En 1867 il met à jour une étude lauréate du Prix Volney afin d'illustrer concrètement sa théorie de la distinction fondamentale entre les langues dotées vs dénuées d'un format morphologique. Le propos de cet article est de montrer en quoi son argumentation dans cette étude est victime de l'obstacle axiologique (cf. François 2014) qui a jeté une ombre sur les prouesses de la linguistique allemande du XIX<sup>e</sup> siècle : il était inconcevable qu'un peuple primitif parle une langue attestant d'un pouvoir d'abstraction et Steintal trouvait dans les langues des peuples mandingues les preuves de leur inaptitude à toute pensée abstraite.

## MOTS-CLÉS

linguistique allemande, XIX<sup>e</sup> siècle, Steintal (Heymann), langues mandingues, format morphologique

## ABSTRACT

In 1848, Steintal outlined a vast epistemological enterprise merging Humboldt's view of the genius of languages (*Sprachgeist*) as inextricably linked to the spirit of the peoples who speak them (*Volksgeist*) and Hegel's

view of the phenomenology of mind progressing dialectically from age to age. In 1867 he updated a Volney Prize-winning essay to illustrate concretely his theory of the fundamental distinction between languages with vs. without morphological pattern. The purpose of this paper is to show how his argument in this study suffers from the axiological obstacle (cf. François 2014) that cast a shadow on the achievements of 19<sup>th</sup> century German linguistics: it was unthinkable that a primitive people would speak a language attesting to a power of abstraction, and Steinthal found in the languages of the Manding peoples evidence of their inability to abstract thought.

#### KEYWORDS

German linguistics, 19<sup>th</sup> century, Steinthal (Heymann), Manding languages, morphological pattern

---

## 1. L'arrière-plan : Steinthal, Lazarus et la *Völkerpsychologie*<sup>1</sup>

Les notices de l'entrée *Völkerpsychologie* (« psychologie ethnique » dans la traduction de Salomon Reinach, 1909) dans les éditions francophone et anglophone de l'encyclopédie en ligne *Wikipedia*<sup>2</sup> imputent son origine à la somme en dix volumes (1900/1920) de Wilhelm Wundt dont le sous-titre précise « Une étude des lois évolutives dans le langage, le mythe et la coutume », et les noms de Heymann Steinthal et de son beau-frère Moritz Lazarus ne sont mentionnés que marginalement. Seule la version allemande<sup>3</sup> de la notice rend justice aux deux fondateurs de la *Zeitschrift für Völkerpsychologie und Sprachwissenschaft* (Revue de psychologie ethnique et de science du langage), dont le premier numéro est paru en 1860 et qu'ils définissaient en termes anthropologiques (comme

---

1. Je remercie Didier Samain pour sa relecture vigilante d'une première version de cet article. Sur l'œuvre linguistique de Heymann Steinthal dans son temps, voir François (2017b) et sur le premier volume, *Die Sprache*, de la *Völkerpsychologie* de Wundt (1900), voir François (2022).

2. <https://fr.wikipedia.org/wiki/V%C3%B6lkerpsychologie>, et <https://en.wikipedia.org/wiki/V%C3%B6lkerpsychologie>, consultées le 23/02/2023.

3. <https://de.wikipedia.org/wiki/V%C3%B6lkerpsychologie>, consultée le 25/02/2023.

le souligne Ivan Kalmar, 1987 : 673), dans leur introduction intitulée « Considérations introductives sur la psychologie ethnique [...] »<sup>4</sup> :

Que l'homme en tant qu'individu doté d'une âme reste donc l'objet de la PSYCHOLOGIE INDIVIDUELLE, telle que se présente la psychologie jusqu'à présent ; mais qu'il s'y ajoute comme prolongement la psychologie de l'homme en société ou de la société humaine, que nous appelons VÖLKERPSYCHOLOGIE [...] La forme de coexistence de l'humanité est sa division en peuples et l'évolution du genre humain est liée à la disparité des peuples.<sup>5</sup> [trad. JF]

Le thème central de la *Völkerpsychologie* est le *Volksgeist* (occasionnellement *Volksseele*, esprit / âme du peuple). Cette notion, accompagnée de celles de *Zeitgeist* (esprit du temps) et de *Volkskultur* (culture du peuple) figure en allemand dans l'œuvre de Johann Gottfried Herder (cf. Moser 1956) qui s'inspire de Gianbattista Vico (1725) et de *L'esprit des lois* de Montesquieu (1748). Chez Wilhelm von Humboldt (cf. Escoubas 1992 : 63), c'est le troisième pôle du triangle conceptuel constitué par le sens de la langue (*Sprachsinn*) et la vision du monde (*Weltanschauung* / *-ansicht*). Dans son dictionnaire de l'allemand paru en 1807, Joachim Heinrich Campe (précepteur des frères Humboldt, qui avait accompagné Wilhelm à Paris dès l'été 1789 pour s'imprégner ensemble de l'esprit de la Révolution française) définit le *Volksgeist* comme « l'esprit, la manière de penser et de ressentir dominante d'un peuple et de la grande masse dans chaque peuple en général » et il cite une sentence du poète Jean-Paul Richter : « L'esprit du peuple et du temps décide et il est à la fois

---

4. *Einleitende Gedanken über Völkerpsychologie als Einladung zu einer Zeitschrift für Völkerpsychologie und Sprachwissenschaft*. C'est l'auteur de cette contribution qui a traduit toutes les citations figurant en français dans le texte et en allemand en note.

5. „Es verbleibe also der Mensch als seelisches Individuum Gegenstand der individuellen Psychologie, wie eine solche die bisherige Psychologie war; es stelle sich aber als Fortsetzung neben sie die Psychologie des menschlichen Menschen oder der menschlichen Gesellschaft, die wir Völkerpsychologie nennen [...] Die Form des Zusammenlebens der Menschheit ist eben ihre Trennung in Völker, und die Entwicklung des Menschengeschlechts ist an die Verschiedenheit der Völker gebunden.“ (Lazarus & Steinthal 1860 : 5)

l'instituteur et le séminaire de celui-ci »<sup>6</sup>. Mais c'est surtout le juriste Friedrich Carl von Savigny qui a donné ses lettres de noblesse au *Volksgeist* en l'introduisant comme pierre angulaire de sa théorie du droit. Dans l'argumentaire<sup>7</sup> de son essai *Tout droit est issu de l'esprit du peuple : Friedrich Carl von Savigny et la science juridique moderne*<sup>8</sup>, Benjamin Lahusen écrit :

À partir des idées du romantisme, il a élaboré un programme visant à surpasser l'ancien droit naturel et à ouvrir la voie à un nouvel ordre juridique. Mais aussi révolutionnaire que fût sa science, il était politiquement hostile aux avancées de son temps. L'autorité scientifique de Savigny continue d'influencer la doctrine juridique jusqu'à aujourd'hui, tout comme son scepticisme à l'égard des changements politiques.

Kalmar (1987 : 680) considère que Wilhelm Wundt, au long des dix tomes de son ambitieuse *Völkerpsychologie*, a en quelque sorte dévoyé la vision de Steinthal et Lazarus :

Effectivement, Wundt a été décrit comme un « biopsychologue évolutionniste » « parmi les premiers, peut-être le premier à intégrer les idées de Darwin dans son propre système ; et tout au long de sa carrière il n'a cessé de mettre ses idées fluctuantes à ce qu'il comprenait comme la position darwinienne » [Citation empruntée à un article de Robert J. Richards en 1980]. Mais Lazarus et Steinthal n'étaient pas des évolutionnistes ethnocentriques. À l'inverse, ils souscrivaient avec enthousiasme à une éthique pluraliste qui annonce Boas et ses contemporains.

Que Wundt ait eu une conception de la *Völkerpsychologie* biaisée par rapport à celle de Steinthal et Lazarus, et que ces derniers, à l'instar de Michel Bréal en France, aient rejeté la linguistique évolutionniste d'August Schleicher (1861) et sans doute un peu plus tard d'Abel Hovelacque (1876) ne fait pas de doute, en revanche le projet philosophique de Steinthal était explicitement de fusionner

---

6. Campe (1807-1811 : 436) : „*der Geist, die herrschende Denk- und Sinnesart eines Volkes und der großen Masse in jedem Volke überhaupt. „Volks- und Zeitgeist entscheidet und ist der Schulmeister und das Schulmeisterseminar zugleich.“* J.P. Richter »

7. <https://www.velbrueck.de/Programm-oxid/Alles-Recht-geht-vom-Volksgeist-aus.html>, consulté le 25/2/2023.

8. *Alles Recht geht vom Volksgeist aus: Friedrich Carl von Savigny und die moderne Rechtswissenschaft* (2019)

la *Phénoménologie de l'esprit* (1807) et la *Philosophie de l'histoire* de Hegel (1837) et la thèse de Humboldt sur la dichotomie fondamentale entre les langues « dotées d'un format », et celles qui en étaient dénuées (cf. François 2014). À la suite de Humboldt, Steinthal conçoit un rapport analogique entre l'évolution sociétale et morale des peuples et l'évolution morphosyntaxique<sup>9</sup> des familles de langue (représentant « l'épanouissement de l'idée de langue ») comme deux éclairages de la dialectique évolutive de l'esprit du monde (*Weltgeist*) imaginée par Hegel.

## 2. La place de l'étude sur les langues « nègres-mandingues » dans l'œuvre linguistique de Heymann Steinthal

Steinthal n'était en rien un africaniste et sa seule monographie consacrée à une famille de langues africaines avait un objectif clairement idéologique, à savoir illustrer sa conception de la psychologie des peuples 'naturels' dont la langue était dépourvue d'un 'format' morphosyntaxique comparable à celui des langues indo-européennes. La thèse sous-jacente était – comme pour Humboldt – que ces langues représentent un stade primitif du « développement de l'idée de langage » caractérisé par l'APERCEPTION CONCRÈTE (*concrete Anschauung*), tandis que les langues de culture « dotées d'un format » (essentiellement indo-européennes et sémitiques) en représentent un stade accompli, celui des REPRÉSENTATIONS (*Vorstellungen*, cf. François 2017a : 177-178). La monographie occupe une double place dans l'œuvre de Steinthal. Il a soumis la première version en 1851 au jury du Prix Volney et l'a remporté, comme plusieurs autres philologues et linguistes allemands, les indo-européanistes Theodor Benfey en 1831 et Max Müller en 1849, et le

---

9. Dans sa *Caractérisation des types principaux de structure des langues* (1860) Steinthal se concentre sur la présence vs absence d'un format morphologique (le classement syntaxique des mots et les flexions qui les affectent diversement en fonction de ce classement), mais dans sa monographie sur les langues mandingues de 1851 [version lauréate du Prix Volney]/1867 [version éditée], faute de matériau morphologique à évaluer, il se rabat sur le format syntaxique et son argumentation s'y révèle beaucoup moins probante.

slaviste Franz von Miklosisch en 1857. Mais c'est la seconde version de 1867 revue à la lumière de sa *Caractéristique* de 1860, offrant un classement typologique d'inspiration psychologique, qui est devenue la référence<sup>10</sup>.

Le point de départ de cet itinéraire intellectuel, c'est en 1836 la publication posthume par Alexander von Humboldt de *L'Introduction à l'œuvre sur le kavi*, la synthèse théorique de son frère. En 1848, Steinthal consacre sa thèse d'habilitation à l'enseignement universitaire à *La linguistique de Wilhelm von Humboldt et la philosophie de Hegel*. L'idée centrale est que la dialectique de Hegel destinée à suivre l'épanouissement de la raison à travers les peuples et au fil des civilisations est une conceptualisation de génie mais fragilisée par manque de fondement empirique, et que la théorie de Humboldt se résume à un « Classement des langues représenté comme progression de l'idée de langage »<sup>11</sup>. La version primée en 1851 de l'étude des langues mandingues se situe un an après l'ouvrage qui porte ce titre et quatre ans avant sa défense du fondement psychologique de la grammaire au détriment des bases logiques de la grammaire philosophique du XVIII<sup>e</sup> siècle, *Grammaire, logique et psychologie – Leurs principes et leurs relations mutuelles*. Et la version définitive de 1867 se situe entre la *Caractérisation des types principaux de structure des langues* parue en 1860 comme révision du *Classement* de 1850 et *L'Introduction à la psychologie et à la linguistique* de 1871 qui connaîtra une seconde édition en 1881 et figurera finalement en 1893 comme premier volet de *L'Abrégé de linguistique*. À cette époque, Steinthal commence à se détourner de la linguistique générale pour se consacrer exclusivement à la partie éthique de son œuvre et il offre à son collègue suisse Franz Misteli de réviser la *Caractérisation* de 1860 comme second volet (celui-ci, fort du consentement de Steinthal va approfondir et rectifier le classement typologique, cf. François 2014).

---

10. La *Polyglotta Africana* de Sigismund Koelle était parue en 1854 et l'apport de données nouvelles sur les langues mandingues a été l'un des facteurs décisifs pour la révision du mémoire primé en 1851.

11. Titre de la première version en 1850 de la synthèse typologique d'inspiration psycholinguistique de Steinthal, qui deviendra en 1860 la *Caractérisation des types principaux de structure des langues*.

L'œuvre de linguistique générale et psycholinguistique de Steinthal est donc un *work in progress*, marqué par deux révisions successives de son classement typologique des langues entre 1850, 1860 et 1893 (par Fr. Misteli) et la révision de l'étude primée des langues mandingues entre 1851 et 1867. Et son objectif, affiché dans la thèse de 1848, de clarifier la pensée linguistique de Humboldt, est réaffirmé dans un second ouvrage d'hommage à Humboldt publié en 1883, *Les œuvres de philosophie du langage de Wilhelm von Humboldt*.

### 3. Les langues avec vs sans « format » chez Humboldt et Steinthal : le primat des flexions

#### 3. 1. Humboldt entre *Volksgeist* et *Sprachidee*

Dans sa synthèse posthume de 1836, Humboldt reprend le classement morphologique des langues, assez consensuel à l'époque (cf. François 2014 : 132-138) en langues à morphologie isolante (corrélées à un *Volksgeist* primitif, ce qui pose un sérieux problème à propos du chinois mandarin), agglutinante (avec des marques grammaticales invariables dans leurs deux formes autonome et liée) et flexionnelle (avec des marques grammaticales liées distinctes de leur contrepartie autonome). Après s'être plongé dans l'étude de la morphologie de plusieurs langues amérindiennes et notamment celle du nahuatl (qu'il appelle « le mexicain »), il introduit une variante des langues agglutinantes, les incorporantes, caractérisée par des unités lexicales hypercomplexes par intégration de plusieurs affixes grammaticaux et relationnels. Henri Dilbermann (2006 : 165) commente ainsi la notion de « forme (interne) de la langue » (*innere Sprachform*) qui génère ce classement quadripartite et correspond au génie propre (*Sprachgeist*) de chaque nation :

D'une part, la forme de la langue explique que chez un peuple donné la pensée s'exprime sous un mode déterminé, selon des règles déterminées. Cette forme est le secret de l'apparition spontanée de la langue comme un tout organisé, cela malgré la multiplicité des locuteurs. Mais, en soi, cette forme spontanée est l'expression du génie de la nation. Bien que l'individu, membre de cette nation, n'ait pas vraiment constitué la

langue – ou, du moins, sa forme –, il s’y retrouve comme s’il en était l’auteur. C’est que cette forme n’est pas pour lui un carcan. Elle s’accorde merveilleusement à sa psychologie, tributaire de celle de la nation. Elle est la condition de la liberté et de l’aisance de son expression. Certes existent des parlers spécifiques à l’intérieur de l’idiome. N’en demeure pas moins que la parole individuelle continue en fait l’œuvre d’abord indivise qu’est la langue.

Steinthal se veut l’interprète privilégié de la pensée linguistique de Humboldt dont il explicite dans sa thèse d’habilitation l’intérêt majeur à l’aide d’une citation de *l’Introduction à l’œuvre sur le kavi* (1836 : xviii) :

L’observation de l’interaction entre d’une part la variété des langues et la répartition des peuples, et d’autre part la formation de la puissance de l’esprit humain comme un processus s’épanouissant peu à peu à différents degrés et dans de nouveaux formats, dans la mesure où ces deux phénomènes sont aptes à s’éclairer mutuellement, c’est ce qui m’occupera dans ces études.<sup>12</sup>

et qu’il commente en ces termes (Steinthal 1848 : 35) :

Ce faisant, Humboldt nous promet les fondements d’une histoire de l’esprit humain du point de vue du langage. Il se peut maintenant que Humboldt lui-même donne des définitions différentes de la philologie, de la linguistique et de la philosophie : cela ne peut pas nous empêcher de reconnaître dans son œuvre un chef d’œuvre réellement philologique selon notre interprétation, et à vrai dire ce n’est pas la prise en considération particulière des langues prises une à une qui constitue le domaine propre de la philologie, mais l’esprit humain se développant par différents degrés dans les langues et les peuples particuliers.<sup>13</sup>

---

12. “Die Betrachtung des Zusammenhangs der Sprachverschiedenheit und Völkervertheilung mit der Erzeugung der menschlichen Geisteskraft, als einer nach und nach in wechselnden Graden und neuen Gestaltungen sich entwickelnden, insofern sich diese beiden Erscheinungen gegenseitig auszuhellen vermögen, ist dasjenige, was mich in diesen Erörterungen beschäftigt wird.“

13. „Somit (Humboldt 1836, S.XVIII) verspricht uns Humboldt die Grundsätze einer Geschichte des menschlichen Geistes vom Gesichtspunkt der Sprache aus. Mag darum nun Humboldt selbst über Philologie, Sprachwissenschaft und Philosophie andere Bestimmungen geben, als wir : das kann uns nicht hindern, in seinem Werke ein nach unserem Sinne echt philologisches Meisterwerk zu erkennen, und zwar nicht die besondere Rücksicht auf die einzelnen Sprachen und Völker ist das eigentliche Philologische, sondern die in den einzelnen Sprachen und Völkern durch verschiedene Stufen hindurch sich entwickelnden menschlichen Geist.“

### 3. 2. Comment Steinthal (1860) caractérise les types principaux de structure des langues

La version révisée en 1860 du classement typologique des langues<sup>14</sup> initialement proposé en 1850 est basée sur l'opposition entre langues dénuées vs dotées de format morphologique. Dans les premières (classes I-IV) la *Sprachidee* (et donc le *Volksggeist*) sont à l'état embryonnaire, dans les secondes (classes V-VIII) ils connaissent un épanouissement plus ou moins marqué. Le classement est caractérisé par une symétrie poussée avec quatre classes subdivisées en langues à morphologie coordonnante vs à matériau variable et des deux côtés celles de ce dernier type se divisent en trois sous-types (Tableau 1).

En dépit de la révérence que Steinthal manifeste à l'égard des thèses de Humboldt, son classement s'en écarte sur plusieurs points :

- a) Quantitativement le classement de Steinthal comporte deux fois plus de classes que celui de Humboldt, car il se fonde sur une hiérarchie de propriétés.
- b) Alors que dans le classement de Humboldt, une seule classe, celle des langues flexionnelles, est dotée d'un format morphologique, dans celui de Steinthal, quatre classes dénuées d'un tel format font face à quatre classes qui en sont dotées.
- c) Le chinois classique, bien que classé comme une langue coordonnante au même titre que les langues de l'Insulinde et de l'Annam, passe dans les langues dotées d'un format, ce qui indique que Steinthal a tiré profit des arguments de Jean-Pierre Abel-Rémusat, professeur de chinois classique au Collège de France, dans son débat épistolaire en 1827 avec Humboldt à propos du caractère supposé primitif de toute langue isolante<sup>15</sup>, et donc aussi du chinois en dépit de la civilisation raffinée à laquelle il s'associe<sup>16</sup>.

14. Sur la version initiale de ce classement, voir François (2017a : 296-299).

15. Voir dans ce volume l'article d'Alice Vittrant.

16. Humboldt avait fini par admettre que le grec classique et le chinois étaient deux langues dotées l'une et l'autre d'un génie propre compatible avec un tel niveau de civilisation (cf. François 2017a : 310-312 et 2018 : 85).

A. LANGUES DENUÉES DE FORMAT MORPHOLOGIQUE			B. LANGUES DOTÉES DE FORMAT MORPHOLOGIQUE			
1. coordon- nantes	2. à matériau variable ⇄ ⇓ ⇨		1. coordon- nantes	2. à matériau variable ⇄ ⇓ ⇨		
	a) redoublement et préfixes	b) ajouts agglutinés aux racines		c) incorporation	a) agglutination lâche des éléments grammaticaux	b) variation interne de la racine
<b>I</b> langues de l'Inde et de l'Annam	<b>II</b> langues polyné- siennes	<b>III</b> langues oura- lo-altaïques	<b>V</b> chinois	<b>VI</b> égyptien	<b>VII</b> langues sémitiques	<b>VIII</b> sanscrit
WH : isolantes	WH : agglutinantes		WH : isolantes	WH : agglutinantes	WH : flexionnelles	

Tableau 1 : Classement des (familles de) langues proposé par Steinthal en 1860 comparé au classement de Humboldt (1837, WH)

d) Quant à l'égyptien, il n'est pas sûr que la distinction entre le caractère rigoureux ou lâche de l'agglutination des éléments grammaticaux suffise à en faire une langue dotée d'un format, à l'inverse des langues ouralo-altaïques qui en seraient dénuées. On soupçonne aisément que le mode de vie nomade et donc l'absence d'urbanisation de la majorité des peuples de cette famille a joué un rôle aussi, voire plus décisif dans ce classement, face au patrimoine artistique exceptionnel de l'Égypte ancienne, que la variété des procédés d'agglutination lexicale. C'est une des manifestations de ce que j'ai appelé ailleurs (2014) l'« obstacle axiologique » de la majeure partie des productions de la linguistique allemande du XIX<sup>e</sup> siècle.

Quatre extraits de la *Classification* (1860) nous permettront de mieux comprendre comment Steinthal s'approprie la distinction humboldtienne entre langues dénuées et langues dotées d'un format morphologique :

Dans la langue, comme activité libre de l'esprit se représentant lui-même, la forme n'est rien d'autre que la quantité et la qualité de ce qui est représenté [...] Si un peuple se représente ses intuitions confusément, c'est-à-dire de manière informe, il n'y a dans sa représentation de lui-même, dans sa langue, que peu de forme, voire pas la moindre [...] La langue n'est formatée que dans la mesure et de la façon dont elle se forme, c'est-à-dire se représente des formes ou donnent forme au contenu.<sup>17</sup>

Le premier extrait frappe par l'abondance des emplois de *Form* (4) et de ses dérivés (le verbe *formen* et l'adjectif *formlos*, « dénué de forme, informe » employé comme adverbe). La dichotomie humboldtienne entre les langues dotées d'une forme interne – qu'on pourrait dire *formatées* – et celle qui en sont dénuées (les langues *informes* ou *non formatées*) est omniprésente et Steinthal reprend à Humboldt la thèse évolutionniste d'un degré de formatage de

---

17. „In der Sprache, der freien Thätigkeit des Geistes sich selbst sich vorzustellen, ist nur so viel und gerade die Form, wie viel und welche vorgestellt wird (...) Stellt sich ein Volk seine Anschauungen unklar, d.h. formlos vor, so ist in seinem Selbstvorstellen, in seiner Sprache, wenig oder gar keine Form [...] Nur in so weit und in der Weise ist die Sprache geformt, als und wie sie sich formt, d.h. Formen oder den Inhalt unter Formen vorstellt.“ (1860 : 103).

l'esprit de chaque peuple, les critères implicites étant assurément le développement de l'écriture, des échanges commerciaux, des centres urbains et de l'organisation de l'état, conformément à la philosophie de l'esprit et de l'histoire de Hegel, sa seconde source majeure.

La diversité des langues dans leurs moyens sonores est donc conditionnée par la diversité des fins qui doivent être atteintes à l'aide de ces formes sonores, c'est-à-dire par la diversité des manières et des formes dans lesquelles les peuples se sont représentés leurs intuitions. Les langues sont aussi diverses que la conscience des divers esprits collectifs.<sup>18</sup>

Les intuitions qui traversent l'esprit collectif de chaque peuple acquièrent un format, en d'autres termes, elles se structurent, notamment selon les catégories de l'individualité, de l'espace, du temps, de la causalité et de la finalité, selon le mode de vie de chaque peuple, son degré d'organisation sociale et spirituelle. Et les différentes catégories grammaticales, les classes de mots, les types de conjonction et de relateurs, les outils de la mise en perspective temporelle, etc., reflètent directement le degré d'affinement de l'esprit collectif :

Sur la base de ce que nous avons découvert, le devoir du classement des langues ne peut être que d'exposer le progrès manifesté à travers les diverses langues, lequel tient au degré de réalisation de l'idée de langage par les divers peuples. De même que les différentes formations de la nature occupent différents degrés d'une voie évolutive, les différentes langues des peuples sont des degrés ou des empreintes des pas que l'humanité a fait franchir à l'idée de langage.<sup>19</sup>

L'idée de langage est à la philosophie du langage de Steinthal ce que l'esprit absolu est à la philosophie de l'esprit de Hegel. Une langue

---

18. „Die Verschiedenheit der Sprachen in ihren lautlichen Mitteln wird demgemäß bedingt durch die Verschiedenheit der Zwecke, welche durch diese lautlichen Formen erreicht werden sollen, d.h. durch die Verschiedenheit der Weisen und Formen, in denen sich die Völker die Anschauungen vorstellten. Die Sprachen sind so verschieden, wie das Bewußtsein der verschiedenen Volksgeister.“ (1860 : 104).

19. „Nach dem, was wir gefunden haben, kann nur dies die Aufgabe der Eintheilung der Sprachen sein, den in den verschiedenen Sprachen sich kundgebenden Fortschritt, in welchem die Völker die Sprachidee verwirklicht haben, darzulegen. Wie die verschiedenen Bildungen der Natur verschiedene Stufen einer Entwicklungsbahn, so sind die verschiedenen Sprachen der Völker Stufen oder Fußstapfen der Sprachidee der Menschheit.“ (1860 : 105).

tend à la 'perfection' quand elle reproduit la variété des mouvements de l'esprit collectif du peuple de la manière la plus adéquate, c'est-à-dire à la fois la plus exacte et la plus économique. Et plus l'esprit du peuple est amené à se représenter une société complexe riche de multiples fonctions et des mythes riches de divinités, plus la langue dans laquelle ses membres échangent leurs jugements et leurs émotions se complexifie.

Ces thèmes réapparaissent sept ans plus tard dans la préface de l'essai sur les langues mandé :

L'étude qui suit n'est pas destinée seulement aux linguistes, mais aussi aux psychologues ; son objectif n'est pas d'enseigner l'usage d'une langue, mais de présenter un état de fait dans son cadre externe et interne ; à proprement parler, elle ne s'adresse pas à la mémoire, mais à l'intelligence ; et comme son objet est une activité psychique, un mouvement de la conscience, il relève de la psychologie ; ce n'est pas un mécanisme phonétique, mais un système de facteurs d'excitation psychique.<sup>20</sup>

C'est une étude relevant de la psychologie des peuples, elle doit donc être lue et évaluée par les psychologues. Le « cadre interne », c'est le degré de structuration de l'esprit collectif. Le « cadre externe », c'est le mode de vie (endogamie vs exogamie, similitude vs distribution des fonctions sociales, autarcie vs activité d'échanges et/ou de conquêtes, etc.). L'échelle d' « excitation psychique » comme facteur de classement des esprits collectifs est symptomatique de l'émergence à l'époque de la psychophysiologie et elle sera centrale dans le classement des peuples et des langues de William Byrne (1885).

---

20. „Die vorliegende Arbeit ist nicht nur für die Sprachforscher [sic], sondern auch für die Psychologen bestimmt ; ihre Absicht [sic] ist nicht, den Gebrauch einer Sprache zu lehren, sondern einen Thatbestand nach seinem äußern und innern Zusammenhange darzustellen ; sie wendet sich nicht eigentlich an das Gedächtniß, sondern an den Verstand ; und weil ihr Gegenstand eine psychische Thätigkeit, Bewegung des Bewußtseins ist, so gehört er in die Psychologie ; er ist nicht ein Mechanismus von Lauten, sondern ein System psychischer Erregungsmittel.“ (1867 : v).

## 4. La paille et la poutre : le raisonnement de Steinthal sur l'absence de « format » dans les langues mandingues

La section 4 se présente sous la forme d'un choix raisonné d'extraits de l'introduction des deux éditions de 1851 et 1867 qui caractérisent l'esprit général de l'étude et des paragraphes consacrés aux conclusions de Steinthal sur les trois points principaux de son argumentation, les §§ 574-5 sur le « manque d'isolation des représentations », les §§ 576sq sur le « manque de condensation des représentations » et les §§ 613sq sur « le manque d'assemblage des propositions ».

### 4. 1. L'exposé introductif

Dans l'introduction à l'édition de référence, celle de 1867, Steinthal intègre un long extrait de l'introduction à l'édition initiale primée en 1851 afin de clarifier l'évolution de son propos au terme de six années marquées par ses publications majeures de 1850, 1855 et 1860 (cf. section 3). Il s'agit pour lui, à la lumière des premières études de linguistique africaine, de mettre à mal l'ambition, propre aux représentants de la grammaire générale ou philosophique, d'élaborer un schéma de corrélation entre les propriétés universelles de la logique (aristotélicienne) et celles de la grammaire dans le sillage de la *Grammaire* de Port-Royal (1660) et jusqu'à la *Sprachlehre* (Traité de la langue, 1801-3) d'August Friedrich Bernhardt. L'heure est aux métaphores biologiques et Steinthal se conforme au goût du temps en comparant la structure des langues à l'organisation physiologique des espèces animales :

Dans le présent écrit, l'auteur cherche à montrer qu'il existe des langues qui sont réellement incommensurables avec le schème catégoriel des grammairiens philosophes et dont la comparaison avec nos langues indo-européennes supérieurement organisées est aussi peu possible quant à leur composition interne, que celles respectivement d'un insecte et d'un mammifère. De même que l'anatomie comparée identifie différents types ou formes primitives d'animaux, de même il appartient à la recherche linguistique comparative de mettre en évidence différents formats internes des langues, qui se ressemblent peu ou pas du tout. On



Un exemple emprunté à la physiologie permettra de saisir cette distinction ; car dans cette discipline on rencontre quelque chose de tout à fait similaire. Les poumons, les branchies, les trachées sont des notions anatomiques comme la flexion, l'agglutination ou l'isolation appartiennent à l'anatomie des langues. En outre, la respiration est une notion générale, une idée abstraite. En revanche, le rétrécissement et l'élargissement de la cage thoracique présents chez les mammifères et les oiseaux, qui, respectivement, aspire l'air dans les poumons ou le fait refluer de ces derniers, ou bien l'absorption de l'eau fraîche à l'aide de la bouche et son rejet par les branchies, comme cela a lieu chez les poissons, ou encore la manière dont les grenouilles avalent l'air, voilà trois notions physiologiques différentes.<sup>23</sup>

Le questionnement de Steinthal n'est nullement obsolète, car il concerne l'interrogation fondamentale, qui continue à agiter les esprits, de l'application des universaux au niveau des besoins de COMMUNICATION, des REPRÉSENTATIONS et/ou des MOYENS alloués par chaque code linguistique pour véhiculer tous les types de représentations et référer à tous les types de procès (cf. Evans & Levinson 2009 ; François 2015, 2018 : 134-137). Et la dichotomie entre les langues distinguant formellement (c'est-à-dire par des indices phonologiques, morphologiques et/ou syntaxiques) les noms des verbes et les langues dont Michel Launey (1994) a désigné la grammaire comme « omniprédicative » (il faisait allusion en priorité aux langues amérindiennes et spécifiquement au nahuatl classique cher à Humboldt) reste à l'ordre du jour (cf. ci-dessous « Nom et verbe du sanscrit font défaut en *soso* ») :

[...] Exactement de la même manière, l'unité interne des langues affirmée jusqu'à présent [par la grammaire générale] est un phénomène qui ne repose nullement sur le format linguistique, mais sur les formats

---

23. „Ein Beispiel aus der Physiologie soll diese Unterscheidung erläutern; denn in dieser findet ganz Aehnliches statt. Lunge, Kiemen, Luftgänge (Tracheen) sind anatomische Begriffe wie Flexion, Agglutination, Isolirung in die Sprach-Anatomie gehören. Athmen ferner ist ein allgemeiner Begriff, eine abstracte Idee. Die bei Säugethieren und Vögeln vorhandene Verengerung und Erweiterung des Brustkastens dagegen, wodurch die Luft aus der Lunge und in die Lunge geführt wird, oder das Aufnehmen des frischen Wassers vermittelt des Mundes und das Heraustreiben desselben durch die Kiemen, wie es bei den Fischen geschieht; oder das Verschlucken der Luft, wie es bei den Fröschen vorkommt : das sind drei verschiedene physiologische Begriffe.“ (1867 : vii).

logiques et métaphysiques du contenu de la pensée, lequel réside pour cette raison en dehors des langues. De même que la grenouille et l'être humain respirent, de même la langue soso et le sanskrit ont des mots pour exprimer l'activité et les substances ; c.à.d. que les catégories métaphysiques sont les mêmes pour le contenu des représentations, comme le processus physiologique l'est pour la respiration ; mais de même que, physiologiquement, la respiration humaine à l'aide de la poitrine fait défaut chez les grenouilles et chez des espèces encore inférieures, de même [les catégories] nom et verbe du sanskrit font défaut en soso.<sup>24</sup>

Comme je l'ai indiqué plus haut, Steinthal a publié trois ouvrages majeurs sur le classement des langues en rapport avec la psychologie des peuples qui les parlent, entre la version primée de son étude des langues mandingues en 1852 et la version définitive publiée en 1867. Dans l'introduction à cette seconde édition il tire les leçons de l'évolution de sa pensée, notamment son souci affiché en 1855 de substituer la « psychologie rationnelle » à la logique classique comme fondement de sa conception de la science du langage (cf. Trautmann-Waller 2006 : 101-123). Et il n'hésite pas à se vanter de son insolence, celle « que réclame la science » :

Alors qu'à l'époque [en 1851] je me limitais à un point de vue strictement descriptif, la révision devait être maintenant entreprise du point de vue supérieur de la psychologie rationnelle [...] Je me rends bien compte que j'ai posé au psychologue plus de tâches qu'il ne peut en assumer. Mais je considérerai cela déjà comme un succès satisfaisant si j'ai pu rendre un tel problème accessible au psychologue.<sup>25</sup>

---

24. „[...] Ganz ebenso ist auch die bisher behauptete innere Einheit der Sprachen eine Erscheinung, die gar nicht mehr auf der sprachlichen Form, sondern auf den logischen und metaphysischen Formen des Denkinhaltes beruht, die eben darum ganz außerhalb der Sprachen liegt. Wie der Frosch und der Mensch athmet, so haben die Soso-Sprache und das Sanskrit Wörter für Thätigkeit und Substanzen: d.h. wie dort der chemische Prozeß, so sind hier die metaphysischen Kategorien für den Inhalt der Vorstellungen dieselben; wie aber physiologisch genommen das menschliche Athmen vermittelt der Brust bei den Fröschen und noch niedrigeren Thieren nicht zu finden ist, so geht auch dem Soso das Nomen und Verbum des Sanskrit ab.“ (1867 : viii).

25. „Stand ich damals noch auf dem rein deskriptiven Standpunkte, so mußte jetzt die Ueberarbeitung von dem höhern, psychologisch rationalen Standpunkte aus unternommen werden. [...] Ich sehe recht wohl, daß ich in diesem Buche dem Psychologen mehr Aufgaben gestellt als gelöst habe. Aber ich würde schon dies für

Que les langues traitées ici se situent à un niveau très bas, c.à.d. qu'elles soient très imparfaites dans leur organisation, c'est ce que je dois asserter avec l'insolence que réclame la science.<sup>26</sup>

Steinthal n'a cependant pas un esprit obtus et il se doute bien qu'affirmer l'imperfection de langues parlées dans des régions d'Afrique colonisées ou en voie de l'être et qui fournissent des esclaves à l'Amérique du nord et centrale, c'est du pain béni pour les partisans de l'esclavagisme, et il cherche à prendre ses distances à l'égard de ces derniers, mais sans empressement excessif. Car il est aussi philosophe et théologien israélite, et il est convaincu que les peuples dotés d'une conscience morale supérieure ont eu la chance de prendre la bonne voie, mais que ce n'est pas un « hasard malheureux » qui a sclérosé la conscience morale des « sauvages » : ils sont simplement victimes de leur place inférieure dans le grand programme de l'évolution mentale et morale de l'espèce humaine.

Ce raisonnement s'apparente à celui d'August Schleicher (1861) à propos de la hiérarchie des types de langue : les peuples qui parlent une langue à morphologie isolante n'ont pas eu l'énergie nécessaire pour transformer des particules autonomes et faiblement catégorisées en affixes et parler ainsi une langue agglutinante, et ceux qui parlent une langue à morphologie agglutinante n'ont pas su s'élever jusqu'au stade supérieur de la hiérarchie qualitative des morphologies grammaticales, celui des langues flexionnelles. La sclérose que Steinthal flétrit ici consiste dans un premier temps à renoncer à penser les catégories grammaticales de manière suffisamment claire pour les exprimer à l'aide de morphèmes destinés à ce seul usage (pour l'effort menant à l'agglutination) et dans un second temps à renoncer à délimiter optimalement la composition des mots (pour l'effort menant à la flexion). Dans les langues à morphologie agglutinante, les affixes restent formellement apparentés aux particules autonomes, si bien que la délimitation du mot reste floue, et dans les langues incorporantes les mots font l'objet d'une composition exubérante qui

---

*einen genügenden Erfolg halten, wenn ich dem Psychologen ein Problem zugänglich gemacht hätte.* (1867 : ix).

26. *„Daß die hier behandelten Sprachen sehr niedrig stehen, d.h. daß sie in ihrer Organisation sehr unvollkommen sind, muß ich mit der Rücksichtslosigkeit aussprechen, welche die Wissenschaft fordert.“* (1867 : xiv).

témoigne d'un défaut d'ajustement à la dimension idéale des unités lexicales qui caractérisent le grec classique, le sanscrit ou le latin. Il ne reste plus aux chanceux qu'à prendre en charge leurs frères « coupables de leur malheur et de les éduquer », ce qui n'est rien d'autre que le projet de l'administration coloniale :

S'il devait m'arriver le malheur que l'un de ces esprits égarés qui défendent l'esclavagisme, fasse référence à mon livre [dans ce sens], cela me peinerait, mais de la même manière qu'un hasard malheureux peut peiner. [...] Le sauvage est un être humain sclérosé de manière bien trop précoce : effectivement le nègre vit dans une sauvagerie non pas débri-dée, mais constitutive, propre à asservir son esprit.<sup>27</sup>

C'est ainsi que procède la science rigoureuse. Et qu'en résulte-t-il sur le plan pratique ? Rien d'autre ne peut en résulter que l'obligation des peuples plus heureux, dotés d'une moralité supérieure, de prendre soin de leurs frères malheureux et coupables de leur malheur et de les éduquer<sup>28</sup>.

#### 4. 2. Le « manque d'isolation des représentations »

Il reste cependant à démontrer par un examen attentif que les langues mandingues sont effectivement sclérosées. Steinthal entend faire cette démonstration en illustrant trois défauts congénitaux de ces langues, à savoir sur le plan psycho-sémantique le manque d'isolation et de concentration des représentations et sur le plan syntaxique le manque d'assemblage des phrases.

Concernant le premier défaut, Steinthal présente le mot dans les langues « civilisées » comme une « monade », en quelque sorte, pour reprendre la métaphore biologique, comme une cellule constituée

27. „Sollte ich das Unglück haben, daß einer jener verkehrten Geister, welche die Sklaverei in Schutz nehmen, sich auf mein Buch berufe : so würde mich das schmerzen, aber nur wie mich ein unglücklicher Zufall schmerzt. (...) Der Wilde ist ein allzufrüh verknöchertes Mensch; namentlich der Neger lebt nicht in bandenlos schweifender, sondern in constituirter Wildheit, unter Institutionen der geistigen Knechtschaft.“ (1867 : xiv).

28. „So richtet die strenge Wissenschaft. Und was folgt daraus für die Praxis ? Nichts Andres kann daraus folgen, als die Pflicht der glücklichern, der sittlichern Völker, sich ihrer unglücklichen, ja schuldigen Brüder anzunehmen und ihnen Bildung mitzutheilen.“ (1867 : xv).

d'un cytoplasme et d'une membrane l'isolant de son environnement. Et il cite (sans mentionner les formes concernées) des exemples qui selon lui vont à l'encontre de cette définition du mot-monade en soso, l'une des langues mandingues en cause. La notion d'« objet interne » (cf. *vivre sa vie, aller son chemin, pleurer des larmes de crocodile*) offre un premier argument à l'encontre de la thèse d'une disparité absolue entre des représentations lexicales « encapsulées » spécifiques aux langues « civilisées » et des représentations aux limites floues typiques des langues « primitives ». Mais l'allemand, la langue maternelle de Steinthal, présente des cas d'incertitude sur la délimitation des « mots » encore plus intrigants.

Deux exemples suffiront à l'illustrer. Pour évoquer la rencontre entre deux personnes en mouvement, l'allemand offre la construction *jemand*<x :nominatif> *geht/kommt jemandem*<y :datif> *ENTGEGEN* et pour évoquer la position de deux personnes assises face à face *jemand*<x :nominatif> *sitzt jemandem*<y :datif> *GEGENÜBER*. Les grammaires classiques de l'allemand distinguent dans de tels cas deux types d'unités grammaticales, les postpositions (très rares) et les préfixes séparables (très fréquents). La consultation de plusieurs dictionnaires révèle une entrée *gegenüber-sitzen* dans la majorité d'entre eux, cependant le vénérable dictionnaire des frères Grimm et le dictionnaire bilingue Leo ne la mentionnent pas, alors que tous ont une entrée *entgegen-gehen/kommen*. Le consensus règne donc parmi les lexicographes sur le statut de préverbe séparable pour *entgegen*, mais pas pour *gegenüber*. À l'infinitif et au participe passé (les deux formes où un préfixe séparable est raccordé au verbe) on peut donc hésiter orthographiquement entre *ich möchte [Ihnen [gegenübersitzen]* (préverbe séparable raccordé) et *ich möchte [Ihnen gegenüber] sitzen* (postposition) pour « j'aimerais être assis en face de vous », ce qui n'est pas le cas pour *sie ist mir [entgegengekommen]* (« elle est venue à ma rencontre »).

Plus clair encore, pour dire que X contrôle la qualité Y de Z<qqc/qqn>, on emploie la construction *jemand*<x:nominatif> *prüft etwas/jemand*<z:accusatif> *auf etwas*<y:accusatif> *hin*. Aucun dictionnaire ne prévoit une entrée verbale *\*hin-prüfen*, alors que le statut de préverbe séparable est largement reconnu pour *hin-* en composition avec d'autres verbes. Il y a donc un consensus entre lexicographes pour

considérer que la structure en cause implique une « circomposition » : *jemand*<x:nominatif> [*prüft etwas/jemand*<z:accusatif> [*auf etwas*<y:accusatif> *hin*]] et orthographiquement le constituant *hin* ne sera pas rattaché au verbe à l'infinitif ou au participe passé : *sie hat ihn darauf hin geprüft / sie will ihn darauf hin prüfen*.

Ce qui est en cause ici, à l'encontre du raisonnement de Steinthal, c'est qu'en allemand les « mots » n'ont pas nécessairement de délimitation rigoureuse : *entgegenkommen* est indiscutablement un « mot » (plus exactement un polylexème comportant un préverbe séparable), *gegenübersitzen* l'est pour la majorité des lexicographes et \**hinprüfen* n'en est pas un (autrement dit, une « membrane cellulaire » sépare le constituant/particule *hin* du verbe *prüfen*). De ce fait, entre l'allemand et le soso il n'y a sur ce point qu'une différence de degré au long d'un continuum allant du mot au syntagme :

§ 574. Seul ce qui se dissocie mutuellement parvient à une existence indépendante. C'est ainsi que le mot doté d'un format des langues flexionnelles vit dans la conscience comme une monade, comme un élément particulier. Il n'en est pas de même dans les langues mandingues. [...]

Le mot pour *hören* (entendre) ne peut pas être employé comme l'expression isolée, absolue de l'activité de l'oreille comme nous le faisons quand nous disons : j'entends, je ne peux pas entendre ou je t'entends. [En soso] on ne peut pas dire simplement : *decken* (couvrir) ou *schließen* (fermer) et *öffnen* (ouvrir), mais *fermer* ou *ouvrir une ouverture*.<sup>29</sup>

Par ailleurs, Steinthal décompose le verbe *kulo<sup>o</sup>n* en soso en *ku* (chose) et *lo<sup>o</sup>n* (savoir) en précisant que la désignation *ku* a été remplacé par *fe* pour désigner une chose et donc que les peuples Soso ne peuvent pas se représenter la notion de *savoir* sans la mention (ici démotivée) d'un objet de savoir :

On ne peut pas dire simplement : *couvrir* ou *fermer* ou *ouvrir*, mais *fermer* ou *ouvrir une ouverture*. En soso on dit *nya* ou *nai tō* “voir avec l'œil” pour *voir* [...] Soso *kulo<sup>o</sup>n* (savoir) semble déjà composé de *ku*

29. „ §574. Nur was sich durch eigenthümliche Bildung von einander absondert, gelangt zu selbständigem Dasein. So lebt nur das geformte Wort der flectirenden Sprachen im Bewußtsein als eine Monas, als ein besonderes Element. Nicht so in den Mande-Sprachen. [...] Das Wort für „hören“ kann nicht als isolirter, absoluter Ausdruck für die Thätigkeit des Ohres gebraucht werden, was wir thun, wenn wir sagen : ich höre, ich kann nicht hören oder ich höre dich.“ (1867 : 220-221).

(chose) et *lo<sup>o</sup>n* (savoir) ; mais c'est sorti de la conscience des Sosos, car ils ont perdu le mot *ku* ; c'est pourquoi ils disent pour *savoir fe kulo<sup>o</sup>n*, c.à.d. qu'ils adjoignent leur mot vivant pour *chose*.<sup>30</sup>

Steinthal considère donc l'emploi absolu d'un verbe comme une preuve de l'aptitude supérieure de l'esprit « civilisé » à se représenter une action indépendamment de son objet, mais les conditions favorables à l'emploi absolu d'un verbe dépendent seulement de l'accessibilité du référent de l'objet absent pour les interlocuteurs dans la situation de communication et si *je* {*crois / pense / sais*} Ø exprime un contenu expérientiel immédiat, l'emploi absolu est inadéquat dans le passé : ? *j'ai* {*cru / pensé / su*} Ø ou dans le futur : ? *je croirai / penserai / saurai*} Ø.

Quant au processus de démotivation lexicale, on le rencontre aussi dans les langues « civilisées », notamment dans des expressions figées comme *au jour d'hui* > *aujourd'hui* où non seulement \**hui* (< lat. *hodie*) mais même *jour* se démotive dans *au jour d'aujourd'hui*. De même la préposition *à* perd progressivement toute réalité dans la série chronologique suivante au point de réapparaître subrepticement au pluriel (*aux*) :

*tour* > *entour* > *à l'entour* > *alentour* > *les alentours* > *aux alentours* > ?*alentourer*  
 n.m. n.m. groupe prép. adverbe groupe nom. groupe prép. verbe<sup>31</sup>

### 4. 3. Le « manque de condensation des représentations »

Steinthal considère que la construction ci-dessous en soso, caractérisée par la subordination à la proposition d'action (qqn roule une pierre) d'une proposition de mouvement consécutif (la pierre s'en va du chemin) « manque de condensation », ce qui revient à dire que les Sosos se représentent l'action et le mouvement

30. „Man kann nicht ohne Weiteres sagen : decken oder schließen oder öffnen, sondern eine Oeffnung schließen oder öffnen (vrgl. § 422). Im Soso sagt man *nya* oder *nai tō* mit dem Auge sehen für *sehen* [...] S. *kulo<sup>o</sup>n* wissen scheint schon zusammengesetzt aus M. *ku chose*, *lo<sup>o</sup>n* wissen; doch dies ist aus dem Bewußtsein der Soso geschwunden, da sie das Wort *ku* verloren haben; darum sagen sie für *wissen* *fe kulo<sup>o</sup>n*, d.h. sie fügen ihr lebendes Wort für *Sache zu*.“ (1867 : 222).

31. Rencontré dans le navigateur Google : « ça doit alentourer les 50€ ».

consécutif comme deux événements dont la relation de cause à effet n'est pas clairement marquée. Mais là encore il faut s'imaginer un continuum de constructions où les relations de temps, de lieu, de cause-conséquence, etc., sont plus ou moins explicitées. En anglais, la conjonction *that* introduisant une complétive tend à être omise quand la relation complétive est aisément accessible (cf. *I suspect Ø you are right*), ce qui n'est pas possible en français (\**Je suppose tu as raison*). Si l'on suit le raisonnement de Steintal, le français qui exige l'explicitation de la relation complétive est donc plus « civilisé » que l'anglais !

<i>Ī</i>	<i>sém-me</i>	<i>kúmbūru</i>	<i>ám bo kírā-ma</i>
toi	Pierre-cette	roule	[qu'] elle s'en aille chemin-hors de
⇒ roule cette pierre hors du chemin <sup>32</sup>			

#### 4. 4. Le « manque d'assemblage des propositions »

C'est là le troisième défaut des langues mandingues stigmatisé par Steintal. Il distingue entre des images élémentaires que le Soso met en phrase tout comme les langues européennes et des images complexes qu'il serait obligé de décomposer, parce que son pouvoir d'abstraction est limité. Mais en quoi consistent ces « complexes » ? La question de la hiérarchie syntaxique entre l'expression du mouvement et celle du déplacement (intra-local vs trans-local, dans la terminologie de Tesnière 1969, chap.131 : 307-310) permet de comparer deux types de « complexes »<sup>33</sup>. Bally (1965, 4<sup>e</sup> éd., § 581 : 350) donne un exemple typique de subordination de l'expression du déplacement à celle du mouvement en allemand : *Wir segelten vom Ufer her über den Fluss hin nach der Insel zu*, avec trois circompositions de lieu d'origine (*von ~ her*), de passage (*über ~ hin*) et de destination (*nach ~ zu*). Malblanc (1968 : 93) reprend cet exemple qu'il propose de traduire soit en reproduisant le génie (ou, dans les termes de Humboldt, la forme

32. „Wälze den Stein fort von der Straße, Ī sé-m-me kúmbūru ám bo kírā-ma, d.h. du Stein-diesen wälze er möge fortkommen Weg-aus.“ (1867: 223, § 574).

33. L'analyse la plus fine et actuellement privilégiée de cette question est la distinction typologique entre langues à cadrage verbal vs satellitaire élaborée par Leonard Talmy, cf. [https://fr.wikipedia.org/wiki/Cadrage\\_verbal\\_ou\\_satellitaire](https://fr.wikipedia.org/wiki/Cadrage_verbal_ou_satellitaire) (page consultée le 23/02/2023).

interne) de l'allemand : *Ils passèrent à la voile de la rive du fleuve dans l'île, soit conformément au génie du français : Ils mirent à la voile, quittèrent la rive, traversèrent le fleuve et abordèrent dans l'île.* Il est symptomatique que Steinthal, si influencé par la notion de génie ou de forme interne des langues, n'ait pas imaginé que la structure de la phrase examinée dans le paragraphe précédent puisse être typique du génie des langues mandingues. La raison est simple :

- [1] Les « nègres mandingues » ne peuvent pas conceptualiser les actions, états et événements de manière complexe, ils sont forcés de les décomposer en situations élémentaires ;
- [2] leur 'aperception' élémentaire [*Anschauung*] des situations les empêche d'en hiérarchiser les composantes ;
- [3] donc ils ne peuvent pas se faire une image abstraite organisée des situations complexes et ils ne peuvent qu'énumérer leurs composantes, CQFD !

Si le nègre veut se représenter deux images, il a besoin pour cela de deux représentations identiques associées couvrant pleinement les complexes. Même si dans ces deux complexes une représentation est commune, elle est conservée dans chacun des deux. À l'inverse, dans notre esprit, les représentations de chaque complexe ne sont pas moins fermement associées, mais simultanément elles sont autonomes au point que chacune peut connaître des destins qui ne touchent pas le complexe en lui-même. Car grâce au format de nos mots, chaque représentation conserve sa pleine indépendance, alors qu'en même temps leur association (c.à.d. leur relation au tout) reste préservée.<sup>34</sup>

---

34. „Will sich der Neger zwei Anschauungen vorstellen, so bedarf er dazu zwei dieselben völlig deckender Complexe associirter Vorstellungen. Ist nun auch in diesen beiden Complexen eine Vorstellung gemeinsam, so wird diese doch in jedem Complexe festgehalten. Dagegen sind in unserem Geiste die Vorstellungen jedes Complexes nicht minder fest associirt, aber zugleich so selbständig, daß jede einzelne gewisse Schicksale erfahren kann, von denen der Complex nicht mit ergriffen wird. Denn durch unsere Wortform erhält jede Vorstellung ihre volle Selbständigkeit, während doch zugleich auch ihre Association (ihre Beziehung zum Ganzen) erhalten bleibt.“ (1867 : 248).

[§ 633] Dans la conscience du nègre Mandé la perception concrète avec son environnement matériel est encore prépondérante et sa conversion en représentation s’effectue de manière incomplète.<sup>35</sup>

La place respective de l’aperception [*Anschauung*] et de la représentation [*Vorstellung*] selon Steinthal se laisse visualiser comme suit (Fig. 1) :

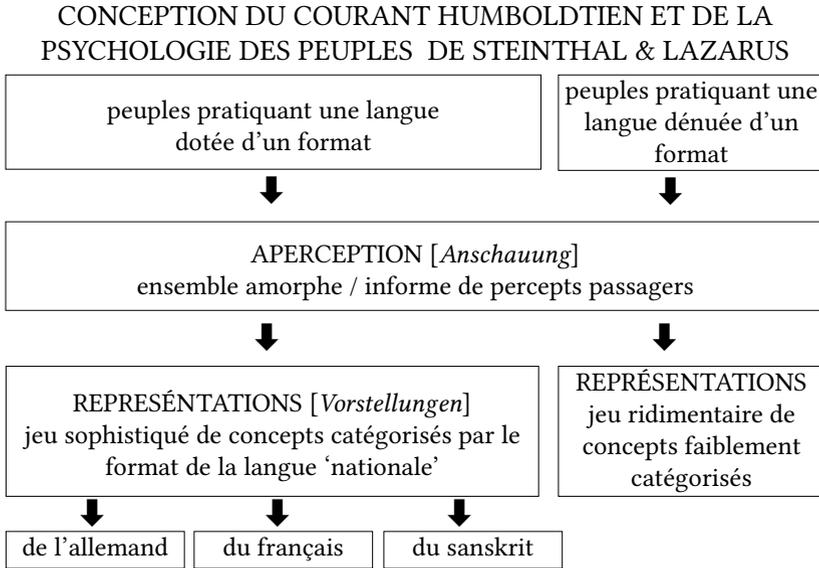


Figure 1 : *Anschauung* / *Vorstellung* dans l’attribution d’un format aux langues

## 5. La linguistique en émergence avait-elle besoin de la psychologie des peuples ?

Relativement aux connaissances et au vocabulaire de l’époque (fournis à Steinthal par le *Polyglotta Africana* de W. Koelle, 1854), l’analyse des données disponibles sur la structure interne des langues

35. „[§ 633] [...] *im Bewußtsein des Mande-Negers ist die concrete Anschauung mit ihren materiellen Verhältnissen noch vorwiegend, und ihre Umsetzung in Vorstellungen ist unvollständig vollzogen.*“ (1867 : 255).

mandingues semble pertinente. Mais le cadre épistémologique est typique des pétitions de principe évolutionnistes du XIX<sup>e</sup> siècle :

1. L'être humain est obligatoirement l'aboutissement de l'évolution phylogénétique de l'espèce humaine et l'homme civilisé est obligatoirement l'aboutissement de l'évolution culturelle des sociétés.
2. L'une des marques du progrès de la culture est la spécification catégorielle des noms et des verbes. Une langue qui ne dispose pas de telles marques (ex. le soso, le nahuatl) est nécessairement inférieure à une langue qui en dispose (ex. le sanskrit).
3. De même une langue qui possède des marques grammaticales de nombre, de personne, de temps (ex. les langues indo-européennes), etc., est supérieure à une langue (ex. le chinois) qui n'en possède pas.
4. C'est la morphologie qui fournit le meilleur classement des langues, avec deux classes superordonnées, les langues dotées vs dénuées d'un format (morphologique) et le classement axiologique des langues et celui des peuples sont parallèles, car la présence de morphèmes grammaticaux affixés à des mots-racines (et notamment si ces morphèmes ont une forme affixale distincte de leur forme indépendante) est interprétée comme un signe de structuration de la morphologie de la langue en accord avec l'organisation supérieure de l'esprit du peuple qui la parle.
5. La syntaxe donne lieu à une analyse plus superficielle, mais il subsiste l'idée de Humboldt que la subordination (notamment préposée) est un acquis substantiel de l'esprit humain civilisé.

Dans son article sur « La *Völkerpsychologie* de Lazarus et Steinthal et le concept moderne de culture » (1987 : 671), Ivan Kalmar rappelle à juste titre :

*In the nineteenth century "culture" was seen as cumulative, developing a single line of human achievement and culminating in the "highest" manifestations of Western civilization. This ethnocentric view, associated with the idea of unilinear social evolution, was gradually replaced by the modern, pluralist view of culture as the distinctive cognitive-behavioral system of unique human groups.*

C'est l'anthropologue et linguiste Franz Boas qui est habituellement considéré comme le catalyseur de ce « changement de paradigme », notamment avec l'introduction à son *Handbook of American Indian Languages* (1911, cf. François 2014 : 147-150), mais Kalmar ajoute curieusement : « *In many respects, Lazarus and Steinthal were thoroughly 'Boasian' before Boas's time* ». L'un des objectifs de Boas et du courant de l'anthropologie culturelle américaine qu'il a fondé, avec notamment les ethnolinguistes Alfred Kroeber et Edward Sapir, était de reconnaître à chaque langue un génie propre, et la distinction de Humboldt et Steinthal entre langues dotées vs dénuées d'une forme interne leur est étrangère. En outre on a vu dans la section précédente que le discours de Steinthal est explicitement paternaliste et justifie l'entreprise colonialiste par ses vertus morales. Au vingtième siècle le *Style de la langue française* (Fritz Strohmeyer, 1910) s'inscrit dans le sillage de la *Völkerpsychologie* et le courant de la stylistique comparée louvoiera dès lors entre deux visions, l'une marquée par l'association entre langue et nation (de Fritz Strohmeyer à Lucien Malblanc, 1968), l'autre proprement linguistique et libérée de l'« obstacle axiologique » qui a partiellement entaché le courant néo-humboldtien, avec Charles Bally (1932), Jean-Pierre Vinay & Jean Darbelnet (1956) et Mario Wandruszka (1969).

## Références

- Bally, Charles. 1965 [1932]. *Linguistique générale et linguistique française* (4<sup>e</sup> éd.). Berne : Francke.
- Benfey, Theodor. 1842. *Griechisches Wurzellexikon*. Berlin : Reimer.
- Boas, Franz. 1911. Introduction. *Handbook of American Indian Languages*, dir par Franz Boas. Washington : Government printing office. 1-84.
- Byrne, William. 1885. *General principles of language*, 2 vol. Londres : Trübner.
- Campe, Joachim Heinrich. 1807-1811. *Wörterbuch der deutschen Sprache*. Braunschweig : Schulbuchhandlung.
- Dilberman, Henri. 2006. Wilhelm von Humboldt et l'invention de la forme de la langue. *Revue philosophique de la France et de l'étranger* 131. 163-191.
- Escoubas, Éliane. 1992. La 'Bildung' et de 'sens de la langue' : Wilhelm von Humboldt. *Littérature* 86. 51-71.

- Evans, Nicholas & Stephen C. Levinson. 2009. The Myth of Language Universals: Language Diversity and Its Importance for Cognitive Science. *Behavioral and Brain Sciences* 32(5). 429-48 ; discussion 448-494
- François, Jacques. 2014. La difficile affirmation de la linguistique générale en Allemagne 1806-1911) et le dépassement de l'obstacle axiologique. *Bulletin de la Société de linguistique de Paris* 109(1). 121-154.
- François, Jacques. 2015. Existe-t-il des universaux du langage ? *Les clés du langage : Nature, origine, apprentissage*, éd. par Jean-François Dortier & Nicolas Journet. Auxerre : Éditions Sciences Humaines. 29-32.
- François, Jacques. 2017a. *Le siècle d'or de la linguistique en Allemagne – de Humboldt à Meyer-Lübke*. Limoges : Lambert-Lucas.
- François, Jacques. 2017b. Trois linguistes allemands du XIX<sup>e</sup> siècle épris de philosophie : Reisig, Steinthal et Schleicher. *De la passion du sens en linguistique - Hommage à Danielle van de Velde*, dir. par Nelly Flaux et al. Valenciennes : Presses Universitaires de Valenciennes. 33-47.
- François, Jacques. 2018. *De la généalogie des langues à la génétique du langage : Une documentation interdisciplinaire raisonnée*. Paris & Louvain : Peeters.
- François, Jacques. 2022. Wilhelm Wundt et la parole au service de l'âme collective. *Revue germanique internationale* 35. 81-96.
- Hegel, Georg Wilhelm Friedrich. 1807. *Phänomenologie des Geistes*. Bamberg & Würzburg : Goebhardt.
- Hegel, Georg Wilhelm Friedrich. 1837. *Vorlesungen über die Philosophie des Weltgeschichte*, hrsg. von Eduard Gans. Berlin : Duncker & Humblot.
- Hovelacque, Abel. 1876. *La linguistique*. Paris : Reinwald.
- Humboldt, Wilhelm von. 1836. *Über die Verschiedenheit des menschlichen Sprachbaues*. Berlin. [trad. fr. par P. Caussat. *Introduction à l'œuvre sur le kavi et autres essais*. Paris : Seuil. 1974].
- Kalmar, Ivan. 1987. The Volkerpsychologie of Lazarus and Steinthal and the Modern Concept of Culture. *Journal of the History of Ideas* 48. 671-690.
- Koelle, Sigismund. 1963 [1854]. *Polyglotta Africana: Or a Comparative Vocabulary of Nearly Three Hundred Words and Phrases in More than Hundred Distinct African Languages*. Graz : Akademische Druck- und Verlags-Anstalt.
- Lahusen, Benjamin. 2019. *Alles Recht geht vom Volksgeist aus: Friedrich Carl von Savigny und die moderne Rechtswissenschaft*. Weilerswist-Metternich : Dittrich Verlag.
- Launey, Michel. 1994. *Une grammaire omniprédicative. Essai sur la morpho-syntaxe du nahuatl classique*. Paris : CNRS-Éditions.

- Lazarus, Moritz & Heymann Steinthal. 1860. Einleitende Gedanken über Völkerpsychologie als Einladung zu einer Zeitschrift für Völkerpsychologie und Sprachwissenschaft. *Zeitschrift für Völkerpsychologie und Sprachwissenschaft* 1. 1-73.
- Malblanc, Lucien. 1968. *Stylistique comparée du français et de l'allemand*. Paris : Didier.
- Moser, Hugo. 1956. Volk, Volksgeist, Volkskultur. *Zeitschrift für Volkskunde* 53.1956/57. 127-140.
- Müller, [Friedrich] Max. 1849. Comparative philology of the Indo-European languages in its bearing on the early civilisation of mankind [Manuscript lauréat du Prix Volney].
- Montesquieu [Charles Louis de Secondat]. 1748. *L'esprit des lois*. Genève.
- Reinach, Salomon. 1909. Compte rendu de Wilhelm Wundt. *Völkerpsychologie. Eine Untersuchung der Entwicklungsgesetze in Sprache, Mythos und Sitte*. Tome III. *Die Kunst. Journal des Savants* 1909(5). 234.
- Richards R.J. 1980. Wundt's early theories of unconscious inference and cognitive evolution in their relation to Darwinian biopsychology. In *Wundt Studies : A centennial collection*, W.G. Bringmann & R.D. Tweney (eds). Toronto.
- Schleicher, August. 1861. *Compendium der vergleichenden Grammatik der indogermanischen Sprachen* (2 vol.). Weimar : Böhlau.
- Schmitt Stéphane. 2001. Type et métamorphose dans la morphologie de Goethe, entre classicisme et romantisme. *Revue d'histoire des sciences* 54(4). 495-521
- Steinthal, Heymann. 1848. *Die Sprachwissenschaft Wilhelm von Humboldt's und die Hegelsche Philosophie*. Berlin [thèse d'habilitation].
- Steinthal, Heymann. 1850. *Die Classification der Sprachen dargestellt als die Entwicklung der sprachidee*. Berlin : Dümmler.
- Steinthal, Heymann. 2004 [1855]. *Grammatik, Logik und Psychologie – Ihre Prinzipien und ihr Verhältnis zu einander*. Berlin :Elibron Classics.
- Steinthal, Heymann. 1860. *Charakteristik der hauptsächlichsten Typen des Sprachbaues* [Caractérisation des types principaux de structure des langues]. Berlin : Dümmler.
- Steinthal, Heymann. 1867. *Die Mande-Neger-Sprachen – psychologisch und phonetisch betrachtet*. Berlin : Dümmler.
- Steinthal, Heymann. 1893 [1871]. *Einleitung in die Psychologie und Sprachwissenschaft*. ( = *Abriß der Sprachwissenschaft*, vol.1). Berlin : Dümmler.
- Steinthal, Heymann. 1883. *Die sprachphilosophischen Werke Wilhelm's von Humboldt*. Berlin : Dümmler.

- Steinthal, Heymann & Franz Misteli. 1893. *Abriss der Sprachwissenschaft* [Abrégé de linguistique]. Berlin : Dümmler.
- Strohmeyer, Fritz. 1910. *Der Stil der französischen Sprache*. Berlin : Weidman.
- Trautmann-Waller Céline. 2006. *Aux origines d'une science allemande de la culture. Linguistique et psychologie des peuples chez Heymann Steinthal*. Paris : CNRS-Éditions.
- Tesnière, Lucien. 1969. *Éléments de syntaxe structurale*. Paris : Klincksieck.
- Vico, Gianbattista. 1725. *Principi di una scienza nuova*. Naples.
- Vinay, Jean-Pierre & Jean Darbelnet. 1956. *Stylistique comparée du français et de l'anglais*. Paris : Didier.
- Wandruszka, Mario. 1969. *Die Sprachen, vergleichbar und unvergleichlich*. Munich : R. Piper Verlag.
- Wundt, Wilhelm. 1900. *Die Sprache*. (= *Die Völkerpsychologie*, vol.1). Leipzig : Engelmann.

# La hiérarchisation des langues vue du Brésil au XIX<sup>e</sup> siècle : le traitement de la question par Júlio Ribeiro (1880)<sup>1</sup>

**José Edicarlos De Aquino**

Universidade Federal do Tocantins

---

## RÉSUMÉ

En 1880, le grammairien Júlio Ribeiro écrit le livre *Traços Geraes de Linguistica*, le premier ouvrage de « linguistique » au Brésil, dans lequel il établit une hiérarchisation des langues lors du traitement de questions sur l'origine et l'évolution du langage et la typologie des langues. L'auteur décrit sept grandes périodes de développement du langage. Cette périodisation est directement alignée sur le développement de l'humanité elle-même et institue une hiérarchie des langues en fonction de leur degré de complexité et de développement. On analyse ici le rapport entre les classements des langues selon des caractéristiques formelles, mais aussi en fonction d'éventuels critères esthétiques, sociologiques et politiques, et d'évaluations des locuteurs, en soulignant la place des langues et des locuteurs du Brésil dans le tableau hiérarchique des langues.

## MOTS-CLEFS

hiérarchisation des langues, origine du langage, évolution du langage, typologie des langues, Brésil, XIX<sup>e</sup> siècle

---

1. Cet article reprend certaines idées présentées dans la thèse de doctorat de l'auteur (De Aquino 2016), intitulée *Júlio Ribeiro na história das ideias linguísticas no Brasil*, soutenue à l'Universidade Estadual de Campinas en co-tutelle avec l'Université Sorbonne Nouvelle en 2016.

**ABSTRACT**

In 1880, the grammarian Júlio Ribeiro wrote the book *Traços Geraes de Linguística*, the first «linguistic» work in Brazil, in which he established a hierarchy of languages when dealing with questions about the origin and evolution of language and the typology of languages. The author described seven major periods in the development of language. This periodization is directly aligned with the development of humanity itself and establishes a hierarchy between languages according to their degree of complexity and development. Here we analyze the relationship between the classifications of languages according to formal characteristics, but also according to aesthetic, sociological and political criteria, as well as to speakers' evaluations, highlighting the place of Brazilian languages and speakers in the hierarchical table of languages.

**KEYWORDS**

hierarchy of languages, origin of language, evolution of language, typology of languages, Brazil, 19<sup>th</sup> century

---

## 1. Introduction

En 1880, Júlio Ribeiro, l'un des plus importants grammairiens brésiliens du XIX<sup>e</sup> siècle, écrit le livre *Traços Geraes de Linguística*, qui a pour particularité d'être le premier ouvrage de linguistique dans l'histoire de la réflexion sur la langue au Brésil.

Outre un avertissement au lecteur et une introduction, l'ouvrage est composé de 9 chapitres qui suivent les thèmes privilégiés de la recherche linguistique au XIX<sup>e</sup> siècle : 1) Siègne du langage articulé ; 2) Appareil de phonation et sons élémentaires ; 3) Origine et développement du langage articulé ; 4) Évolution linguistique ; 5) Langues monosyllabiques ; 6) Langues agglutinantes ; 7) Langues amalgamantes ; 8) Langues contractées ; 9) La science du langage.

Cette œuvre met fortement en valeur une hiérarchie des langues en traitant des questions de l'origine et de l'évolution du langage, ainsi que de la typologie des langues. Le traitement de ces questions par Júlio Ribeiro est précisément l'objet de l'analyse de cet article, qui porte moins sur le tableau et les divisions des langues au Brésil que sur les idées d'un auteur brésilien sur le développement et la classification des langues à la fin du XIX<sup>e</sup> siècle.

## 2. L'origine du langage : développement des langues en parallèle au développement de l'espèce humaine

En traitant de l'origine des langues, l'un des problèmes les plus controversés de l'histoire des sciences du langage, Júlio Ribeiro rejoint les groupes d'études qui se sont développés de façon spectaculaire au XIX<sup>e</sup> siècle dans ce domaine, malgré l'interdiction de la *Société de linguistique de Paris* sur le sujet, et qui ont été dirigés, comme le rappelle Auroux (2008), par les grands noms de la linguistique de l'époque, tels qu'August Schleicher, Max Müller, Abel Hovelacque et André Lefèvre.

En Allemagne, Schleicher, selon Desmet (1996), rejette l'idée du langage comme invention de l'homme ou don divin et récuse l'hypothèse d'une origine commune des langues, arguant que le développement du langage va au même rythme que le développement du cerveau et des organes de la parole. Max Müller (1855), à son tour, se met à caractériser péjorativement les théories de l'origine du langage, en expliquant qu'il est impossible de résoudre historiquement le problème de l'origine du langage et en affirmant, à l'instar de Schleicher, que la science du langage ne peut aborder le sujet qu'à partir du moment où les langues, une fois constituées, commencent à se développer dans l'histoire.

En France, Hovelacque (1877) considère l'origine du langage comme un problème de linguistique propre et légitime, tout en rejoignant Schleicher dans la lutte contre l'idée d'une origine divine des langues. Hovelacque soutient que le problème de l'origine du langage peut être résolu en introduisant dans la linguistique la doctrine du transformisme basée sur la variabilité des espèces. Toujours en France, Lefèvre (1877) répète les protestations de Schleicher, Müller et Hovelacque contre la croyance de l'origine divine des langues et de l'hypothèse du monogénétisme linguistique, en faisant valoir que l'embryon, où le germe du langage est constitué par le cri, est commun aux animaux et à l'homme, établissant ainsi trois époques préhistoriques (interjectif, démonstratif, attributif) et quatre périodes historiques pour le développement du langage (monosyllabique, agglutinatif, amalgamant, contracté).

L'établissement d'une division entre une période historique et une période préhistorique pour le développement du langage articulé est loin d'être un expédient inhabituel à l'époque. Utilisée par des auteurs comme Schleicher et Hovelacque, une telle division sert souvent à promouvoir l'idée d'un processus de maturation et de décadence ultérieure des langues. À ce stade, Klippi (2010) nous montre que si Schlegel interprète l'évolution linguistique en termes de dégénérescence, des noms comme Humboldt et Grimm ont pris les transformations linguistiques comme marques de progrès.

Júlio Ribeiro, traitant de la question de l'origine des langues à partir de Lefèvre (1877), et se plaçant ainsi dans une filiation française qui lie intrinsèquement la linguistique et l'anthropologie, reliera le développement des langues au développement de l'espèce humaine elle-même, et au développement et à la hiérarchisation des races.

La première période de développement du langage articulé, l'interjectif, aurait pour protagoniste un anthropoïde, « grand-père de l'humanité », doté d'un cerveau propice au langage. Ce cerveau aurait permis à l'appareil phonique de « traduire » les différentes impressions en différents cris. Dans ce processus, l'exercice de la voix aurait renforcé et perfectionné l'appareil phonique lui-même de sorte qu'à travers les générations, la pensée et le langage articulé se sont réunis :

*Isto posto, concebe-se o anthropoide avô da humanidade como dotado de um cerebro impressionabilissimo e relativamente perfeito. A terceira circumvolução frontal desse organ, desenvolvida pela lei fatal da evolução, habitava-o a incumbir o seu bem conformado aparelho phonico de traduzir gritos varios as suas diferentes impressões.*

*Traduziu-as e, reflectindo sobre essa conquista da expressão vocal, reteve-a, ampliou-a, deu-lhe variedade, creou a linguagem articulada.*

*Pelo exercicio diuturno da voz reforçou-se e aperfeiçoou-se o aparelho phonico, e a faculdade adquirida foi-se transmittindo de geração em geração pela hereditariedade, polindo-se cada vez mais.*

*Com o correr do tempo as concepções homologaram-se com os seus instrumentos, as idéias hyposthatizaram-se com as palavras, e o pensamento e a linguagem articulada começaram a marchar par a par, auxiliando-se mutuamente, sem mais se poderem separar. [...]*

*O anthropoide antecessor da humanidade, topando com uma pedra, magoando-se, deu um grito, voz simples ou reforçada por uma aspiração,*

*como as nossas interjeições ah ! ih ! oh ! Ao deparar-se-lhe uma arvore carregada de fructos já conhecidos, a alegria irrompeu-lhe do larynge em fôrma de exclamação de jubilo.*

*Depois, ao entristecer-se ou alegrar-se elle sem causa physica visivel, pela reacção interna do systema nervoso sobre o cerebro, moralmente emfim, as mesmas manifestações phonicas reproduziram-se, fundadas na analogia das impressões moraes com as physicas.* (Ribeiro 1880 : 32-34)

‘Cela dit, l’anthropoïde grand-père de l’humanité est conçu comme doté d’un cerveau « impressionnabilissime » et relativement parfait. La troisième circonvolution frontale de cet organe, développée par la loi fatale de l’évolution, lui a permis d’ordonner à son appareil phonétique bien formé de traduire divers cris en différentes impressions.

Il les traduit et, réfléchissant à cette conquête de l’expression vocale, la conserva, l’agrandit, lui donna de la variété, créant le langage articulé.

Par l’exercice diurne de la voix, l’appareil phonétique a été renforcé et perfectionné, et la faculté acquise a été transmise de génération en génération par l’hérédité, et a été de plus en plus raffinée.

Au fil du temps, les conceptions ont été homologuées avec leurs instruments, les idées hypostasiées avec des mots, et la pensée et le langage articulé ont commencé à marcher côte à côte, en s’aidant, mais sans pouvoir être séparés [...]

L’anthropoïde ancêtre de l’humanité, se cognant dans un rocher et se faisant mal, a poussé un cri, d’une voix simple ou renforcée par une aspiration, comme nos interjections *ah! oh! oh!* A la vue d’un arbre chargé de fruits familiers, la joie jaillit du larynx sous la forme d’une exclamation de joie.

Puis, qu’il pleure ou se réjouisse sans cause physique visible, par la réaction interne du système nerveux sur le cerveau, moralement finalement, les mêmes manifestations phonétiques se sont reproduites, fondées sur l’analogie des impressions morales et physiques.<sup>2</sup>

En continuité de ce processus s’est développée la période démonstrative du langage articulé, pendant laquelle le langage apparaît comme un moyen pointu de communication sociale :

*Mais tarde o anthropoide já homem (porque homem foi desde que guardou na memoria a relação que estabelecêra entre a ideia adquirida e o som com que a exprimira), levado pelo viver social a ser util aos seus congeneres, indicou-lhes por gestos acompanhados de gritos, definitivamente articula-*

---

2. Toutes les traductions sont de l’auteur de cet article.

*dos e significativos, o que a experiencia lhe ensinára ser util ou damnoso. Isto, isso, aquillo, aqui, alli, lá, disse elle em seu tosco mas já accentuado meio de communicacão. (Ribeiro 1880 : 34)*

‘Plus tard, l’anthropoïde déjà homme (parce qu’il était homme depuis qu’il gardait en mémoire la relation qui s’était établie entre l’idée acquise et le son avec lequel il l’exprimait), incité par la vie sociale à être utile à ses congénères, indiquant par des gestes accompagnés de cris, clairement articulés et significatifs, dont l’expérience lui apprendra qu’ils sont utiles ou désastreux. *Ceci, cela, ici, là, là-bas* disait-il avec ses moyens de communication grossiers mais déjà structurés.’

Dans la troisième période de développement du langage articulé, appelée période attributive, l’homme, déjà capable de faire des abstractions, est capable de créer des mots qui qualifient les choses elles-mêmes :

*Afinal o cerebro humano robustecido pelo exercido do pensamento começou a fazer abstracções: da pedra que o ferira, da agua que o desalterára, da tarde sombria que o tornára melankholito separou o homem as qualidades que o tinham affectado, e attribuiu-as pela generalisação a todos os objectos ou phenomenos semelhantes. O duro! O fresco ! O triste ! exclamou elle ao vêr uma pedra em seu caminho, ao ouvir o marmurio de um regato, ao attentar no pôr do sol. (Ribeiro 1880 : 35)*

‘Enfin, le cerveau humain, renforcé par l’exercice de la pensée, comença à faire des abstractions : de la pierre qui le frapperait, de l’eau qui le transformerait, de l’après-midi sombre qui le rendrait mélancolique, l’homme sépara les qualités qui l’avaient affecté et, par la généralisation, les attribua à tous les objets ou phénomènes similaires. *Le dur ! Le frais ! Le triste !* s’exclama-t-il en voyant une pierre sur son chemin, en entendant le murmure d’un ruisseau et en regardant le coucher du soleil.’

Dans la première période historique du développement du langage articulé, les mots monosyllabiques, interjectifs, démonstratifs et attributifs se seraient substantivés et transformés en verbes, exprimant les principales gradations de la pensée, à travers des mots encore sans trace de flexion :

*As palavras interjectivas, demonstrativas e attributivas substantivaram-se, converteram-se em verbos, prestaram-se a exprimir todas principaes gradações do pensamento: o homem poudo externar tudo o que sentia. Mas nessa linguagem, já relativamente perfeita, não apparece ainda traço de flexão; nada de indicacão de genero, de numero, de pessoa; nada de deter-*

*minação de tempo ou pessoa; nada de elementos de relação; nada de preposições; nada de conjunções: cada palavra-raiz, ou antes cada raiz-palavra designa uma ideia cujo valor, cujo sentido preciso é determinado pela sua posição na phrase. (Ribeiro 1880 : 35-36)*

‘Les mots interjectifs, démonstratifs et attributifs étaient substantivés, transformés en verbes, capables d’exprimer toutes les grandes gradations de la pensée : l’homme pouvait exprimer tout ce qu’il ressentait. Mais dans ce langage déjà relativement parfait aucune trace de flexion n’apparaît encore ; aucune indication de genre, nombre, personne ; aucune détermination de temps ou de personne ; aucun élément de relation ; pas de préposition ; pas de conjonction : chaque mot racine, ou plutôt chaque racine mot, désigne une idée dont la valeur, dont le sens précis, est déterminé par sa position dans la phrase.’

Selon Júlio Ribeiro, le chinois, l’annamite, le siamois, le birman, le tibétain, le pégu et le kassia conservent toujours les caractéristiques de la période monosyllabique. Dans la période agglutinative, s’est développé un mode d’expression plus synthétique, dans lequel un radical invariable exprimant une idée principale rejoint différents affixes, chacun apportant des modifications à l’idée principale. Selon le jugement de l’auteur, ce serait un moyen plus rationnel d’expression du langage articulé :

*A esta cruda fôrma da manifestação do pensamento no cyclo historico da linguagem articulada, succedeu um modo de expressão mais synthetico, mais racional, mais commodo: a um radical sempre invariavel que exprimia a ideia principal juntaram-se, para modificá-la, affixos apenas alterados, cuja significação propria só relativamente se conservou. (Ribeiro 1880 : 36)*

‘À cette forme grossière de la manifestation de la pensée dans le cycle historique du langage articulé a succédé un mode d’expression plus synthétique, plus rationnel, plus confortable : à un radical toujours invariable exprimant l’idée principale se sont adjoints, pour le modifier, des affixes seulement altérés, dont la signification n’a été que relativement conservée.’

La période amalgamante du langage articulé est représentée par l’émergence du groupe des langues aryennes ou indo-européennes et des langues sémitiques. C’est durant cette période que se développe la flexion des mots :

*A este periodo de invariabilidade rígida da palavras-raiz seguiu-se a flexão ou possibilidade de modificação da propria fôrma dessa palavra-raiz para exprimir as suas relações com outras. Operou-se uma fusão íntima de radicaes attributivos com raizes demonstrativas atrophiadas em desinencias. Entrevê-se ainda em cada palavra deste periodo linguistico o sentido dos elementos que a compõem; esses elementos são, porém, inseparaveis, formam um todo indivisivel cujas partes concorrem todas para accentuar-lhe a significação propria e particular. (Ribeiro 1880 : 36-37)*

‘Cette période d’invariabilité rigide des mots-racines a été suivie par la flexion ou la possibilité de modification de la forme même de ce mot-racine pour exprimer ses relations avec les autres. Une fusion intime d’un radical attributif avec une racine démonstrative atrophiée en terminaisons a été opérée. Dans chaque mot de cette période linguistique, on voit le sens des éléments qui le composent ; mais ces éléments sont inséparables, ils forment un tout indivisible dont toutes les parties concourent à accentuer sa propre signification particulière.’

La période contractée du langage articulé représenterait, selon les mots de Júlio Ribeiro, la période actuelle du monde civilisé. Cette période aurait duré plus de deux mille ans et créé les modifications qui ont donné naissance aux langues romanes à partir du latin, au romaïc à partir du grec, en plus de la profonde transformation des langues germaniques :

*Mas não para na flexão o progredir da linguagem.*

*As syllabas breves que precediam ou seguiam o accento tonico nas palavras flexionaes atrophiaram-se ou desapareceram; as desinencias causaes obscureceram-se, tornaram-se indistinctas, e as gradações do pensamento que ellas exprimiam foram traduzidas por preposições. (Ribeiro 1880 : 37)*

‘Mais la progression du langage ne s’arrête pas avec la flexion.

Les syllabes brèves qui ont précédé ou suivi l’accent tonique dans les mots flexionnels se sont atrophiées ou ont disparu ; les terminaisons finales ont été obscurcies, rendues indistinctes, et les gradations de pensée qu’elles ont exprimées ont été traduites par des prépositions.’

En retirant la question de l’origine des langues du champ théologique et en la faisant rentrer dans ce qui serait le domaine de l’anthropologie et de la linguistique du XIX<sup>e</sup> siècle, Júlio Ribeiro conçoit un processus de développement du langage qui a le mérite de reconnaître le langage articulé comme la qualité fondamentale et déterminante de l’humanité. Selon Auroux (2008), dans l’Occident

chrétien, la représentation de l'origine des langues est initialement organisée à partir du mythe biblique, et il est nécessaire d'attendre jusqu'au XVIII<sup>e</sup> siècle pour que le thème soit laïcisé par les Lumières et que des auteurs tels que Condillac, Herder, Maupertuis, Rousseau, De Brosses, Court de Gébelin, Monboddo, Hervás et Panduro, Pallas, Adelung et Vater commencent à examiner les processus communs et stables de ce qu'on appelle la « nature » dans un débat qui inaugurerait une révolution philosophique séparant nature et culture, et dans laquelle l'arbitraire de la langue joue le rôle de principe de démarcation. Cependant, selon Auroux (2008), ce n'est qu'au XIX<sup>e</sup> siècle que la question de l'origine des langues prend une orientation naturaliste, considérant la langue comme une espèce naturelle qui naît, grandit et meurt, ce qui a à voir avec la découverte et la diffusion de théories évolutionnistes et les avancées de la grammaire comparée permettant de reconstruire les langues parlées par les ancêtres et dont il ne reste pas de témoignage. Ainsi, Júlio Ribeiro rejoint les nombreux auteurs qui, incités par les développements de la grammaire comparative, en viennent, selon Ernest Renan (1858 : 4-5), « à établir la possibilité de résoudre un tel problème d'une manière scientifique ».

La périodisation du développement du langage présentée par Júlio Ribeiro, à la suite des travaux de Lefèvre, est directement alignée sur le développement de l'humanité elle-même et établit une hiérarchie des langues en fonction de leur degré de développement. Le « progrès du langage », dans l'expression de Júlio Ribeiro lui-même, prend son essor à partir de « cris » et de la « voix simple », passe par un « mode de communication rude mais déjà accentué », par des « gestes accompagnés de cris », atteint des « abstractions » et la « généralisation » à travers un « cerveau renforcé par l'exercice de la pensée », arrive au niveau de « l'expression de toutes les gradations de la pensée » avec un langage « déjà relativement parfait », progresse avec un mode d'expression « plus synthétique, plus rationnel, plus confortable », dépasse le stade de « l'invariabilité rigide du mot-racine » avec la création de la flexion, parvient enfin à la « période actuelle du monde civilisé », lorsque « syllabes », « accent tonique » et « terminaisons » « se sont atrophiés », « ont

disparu » et « se sont assombris », entraînant la modification des langues germaniques et la conversion du latin en langues romanes.

Dans une ligne évolutive ascendante, la langue, qui « s'est transmise de génération en génération par l'hérédité, en se polissant de plus en plus », se déploie, au début, en langues asiatiques, telles que le chinois, l'annamite, le siamois, le birman, le tibétain, le pégou et le kassia, et, finalement, dans les langues romanes. Les langues qui sont au même degré de progression, et donc avec des locuteurs au même stade de développement du cerveau et de l'expression d'idées, seront réunies dans le même groupe, dans la même famille.

### 3. Typologie des langues : famille de langues et hiérarchie des races

La typologie des langues est un élément du comparatisme qui sera affiné tout au long du XIX<sup>e</sup> siècle, mais dont les principes et les jugements de valeur ont peu évolué depuis le modèle établi par Schlegel, comme l'explique Auroux (2000a). Friedrich Schlegel, comme l'enseigne Schmitter (2000), établit une classification des principaux types de langues selon leur « structure interne », en les divisant en langues organiques (flexionnelles), celles qui indiqueraient leurs significations accessoires par un changement de voyelle radicale, et en langues mécaniques (isolantes et agglutinantes), celles qui indiqueraient ces significations accessoires en ajoutant des éléments au radical. Cette typologie des langues en isolantes, agglutinantes et flexionnelles sera reprise par des noms célèbres comme Schleicher, comme le montrent Auroux, Bernard et Boulle (2000). La paternité de cette typologie, pour ainsi dire, était à l'époque un sujet de controverse, comme l'explique Trabant (2000).

Júlio Ribeiro organise les langues en quatre familles, monosyllabiques, agglutinantes, flexionnelles et contractées, reliant chacune, avec un degré plus ou moins détaillé, à une période historique de développement du langage, à un niveau correspondant d'idées, à une région géographique et à un peuple, une race. Dans cette entreprise, il se réfère à Philipp von Martius et Max Müller, travaillant la question principalement à partir de Hovelacque (1877)

et Lefèvre (1877), et pour sa part, il intègre à cette étude la langue portugaise et les langues indigènes brésiliennes.

Les langues monosyllabiques, comme nous l'avons vu, sont représentées par l'auteur par sept langues asiatiques : chinois, annamite, siamois, birman, tibétain, pégou et kassia. Selon Júlio Ribeiro, « les domaines de ces langues ne sortent pas de l'Asie de l'Est : elles sont toutes parlées par des peuples de race jaune » (Ribeiro 1880 : 66). Sur une échelle, cette famille de langues se positionne comme « la plus simple », « la forme élémentaire », comme un « premier état », comme une « première couche linguistique » ou comme un « premier atterrissage historique », car, à ce stade, « les mots ne sont que des racines », c'est-à-dire « la forme du mot est unique : c'est la racine dans sa crudité, c'est la racine invariable » (*ibid.* : 63). Les langues monosyllabiques sont décrites comme présentant un manque car elles exprimeraient les idées d'une manière floue : « Aucune indication sur la personne, le sexe ou le nombre ; aucune indication de temps ou de mode ; pas d'éléments de relation, pas de conjonctions, pas de prépositions. Juste une idée très vaste, sinon très vague, qui ne traduit même pas la forme, déjà si peu déterminée, de notre infinitif » (*ibid.* : 63). « Seules des idées essentiellement générales suscitent ces racines » dans les langues monosyllabiques, pour reprendre les mots de Júlio Ribeiro (*ibid.* : 63). Júlio Ribeiro souligne le rôle de la « position du vocable dans la phrase » et des inflexions de la voix » dans les langues monosyllabiques, (*ibid.* : 64-65), le chinois étant donné en exemple.

Les langues agglutinantes correspondraient à une « seconde période historique » quand « de nombreux éléments se juxtaposent, s'agglomèrent, s'agglutinent réellement » (*ibid.* : 67). Ainsi, dans la question des mécanismes de construction du sens et d'expression des idées, les « éléments qui entrent dans la fabrication du mot n'ont plus leur valeur propre, leur valeur première », de sorte qu'un de ces éléments, un seul, « comporte l'idée principale, la signification, le sens » tandis que les autres « perdraient leur valeur indépendante, tout en conservant toujours une portée personnelle, individuelle, mais tout à fait relative ». L'exemple donné pour « mieux comprendre la Théorie » est celui de la langue Tupi.

Les langues agglutinatives sont, selon les termes de l'auteur, « plus nombreuses aujourd'hui que toutes les autres, et parlées dans les cinq parties du monde » (*ibid.* : 69), divisées en douze classes : les dravidiennes, « parlées par la population de la race jaune, descendante de celle que les Aryens ont trouvée dans la péninsule cisgangaétique », « servant de lien avec environ quarante-cinq millions de personnes des montagnes Vindhya et du ruisseau Narmadâ jusqu'au cap Camorim » (*ibid.* : 69) ; les Mongoles, qui « ont régné en Mongolie, en Russie sur la rive gauche de la mer Caspienne, le Tibet, la Perse et la Sibérie » (*ibid.* : 70) ; tartares ou turques, « éparpillés dans toute l'Asie, des rives de la Méditerranée à la Léna en Sibérie » (*ibid.* : 70) ; samoyèdes, qui « prédominent au bord de la mer glaciaire » (*ibid.* : 71) ; finnoises, « parlées en Hongrie, en Russie et dans certaines régions septentrionales de la Suède et de la Norvège » (*ibid.* : 71) ; boréales, « parlées dans les régions arctiques » (*ibid.* : 72) ; langues du Caucase, désignant génériquement par leur nom la région où elles sont parlées ; africaines, avec indication spécifique d'une région pour une seule langue, « la langue de l'île de Fernando Pó » dans le groupe bantou (*ibid.* : 73) ; océaniques, « parlées en Océanie et en Asie du Sud-Est, en particulier dans la péninsule de Malacca » (*ibid.* : 74) ; américaines, « Langues indigènes d'Amérique » (*ibid.* : 75) ; basque, « parlée sur une zone d'environ 190 kilomètres de large et de 50 à 80 de long » entre la France et l'Espagne (*ibid.* : 77) ; japonais et coréen, « langue du Japon » et « langue de Corée », respectivement (*ibid.* : 79).

Composant un tableau géant qui donne une idée du degré de connaissance et de description des langues parlées dans le monde au XIX<sup>e</sup> siècle, ces douze grandes classes seront subdivisées, pour certaines d'entre elles, en d'autres classes, branches, rameaux, groupes et dialectes, selon la terminologie utilisée dans le texte, manière de concevoir et nommer les langues selon une hiérarchie. Ces subdivisions comportent parfois d'autres subdivisions. C'est le cas des langues américaines, qui auraient « vingt-six groupes d'idiomes distincts mais apparentés », à leur tour subdivisés en « une multitude de rameaux peu connus » (*ibid.* : 75). C'est dans ce groupe de langues américaines que nous trouverons les langues parlées au

Brésil, à partir d'une référence de Júlio Ribeiro à la classification de Carl Friedrich Philipp von Martius.

Certaines langues occupent une place importante dans la famille des langues agglutinantes, étant hiérarchisées du fait de leur statut de langue officielle ou nationale, qu'elles produisent de la littérature et de leur difficulté à être parlées ou analysées. Ainsi, parmi les langues turques, l'Osmanli est « la langue officielle et littéraire de Constantinople » (*ibid.* : 70) ; le magyar, « langue nationale hongroise », « s'est affiné en Europe et s'est illustré par la production d'une littérature remarquable, dont le premier monument date du XII<sup>e</sup> siècle », étant « l'une des langues européennes les plus intéressantes pour la science parce que, malgré la civilisation du peuple qui la parle et bien que vivant étroitement avec les langues de flexion, il n'a jamais dépassé l'état d'agglutination », un fait qui le conduirait à « servir d'exemple de cristallisation linguistique » (*ibid.* : 72), dans une formulation qui traduit la surprise de l'auteur devant l'inadéquation entre le degré de civilisation d'un peuple et le stade de développement d'une langue ; et la langue basque est une « langue curieuse et difficile, dont l'existence en Europe a cédé la place aux divagations des ethnographes et des linguistes » (*ibid.* : 77), une langue « compliquée et obstruée » qui « offre des difficultés insurmontables à ceux qui sont seulement habitués à commenter les textes grecs et latins par les processus empiriques » (*ibid.* : 78).

Correspondant à la troisième et avant-dernière période historique de développement du langage – la période amalgamante – les langues regroupées dans la famille des langues à flexion auraient la propriété « d'exprimer les relations que les mots entretiennent entre eux, non seulement par des affixes, mais aussi par des variations de la forme racine elle-même » (*ibid.* : 80). Ses locuteurs compteraient donc sur la faculté de flexion, à savoir « la possibilité qu'a une racine d'exprimer, en se modifiant, un certain changement de sens » (*ibid.* : 80), une fonctionnalité exprimant des relations de sens des éléments qui composent le mot, indisponible dans les langues monosyllabiques et agglutinantes. Il s'agit, pour l'auteur, d'un « fait tiré des langues indo-européennes » (*ibid.* : 82).

La famille des langues flexionnelles est divisée en trois groupes, le chamitique, le sémitique et l'indo-européen, chacun également

divisé en plusieurs autres sous-groupes et branches. Les hiérarchies sont établies non seulement entre le statut de groupe, de sous-groupe et de branche, mais également entre les trois groupes eux-mêmes, Júlio Ribeiro déclarant, par exemple, que « le groupe chamitique semble représenter le passé des langues sémitiques après l'état d'agglutination » (*ibid.* : 86).

Le portugais apparaîtra dans le groupe des « langues vivantes » du sous-groupe italique du groupe indo-européen. Presque sans indication des peuples et des régions, comme si ce fait était évident, certaines langues de la famille des langues flexionnelles ont pour attribut d'être « mortes » ou « vivantes », mais aussi « anciennes » ou « modernes », « divisées en dialectes », « non classifiées », de la branche « basse », « moyenne » ou « élevée », « liturgique » et ou « ecclésiastique », prédicats qui, une fois de plus, construisent une hiérarchie des langues.

Les « langues indo-européennes vivantes, en particulier les langues romanes et germaniques », considérées comme « langues modernes », au dernier stage du développement linguistique, la « période linguistique actuelle », forment la famille des langues contractées, caractérisées principalement « par l'atrophie ou la disparition des syllabes qui suivent la syllabe accentuée » (*ibid.* : 91). Le texte met en évidence le français, « la langue romane la plus originale et la plus contractée » (*ibid.* : 90), et le portugais, « la langue romane qui a le plus fidèlement conservé les formes de la conjugaison latine » (*ibid.* : 96), dans un discours récurrent depuis la Renaissance cherchant à lier les langues au latin et donc à les valoriser.

Dans une posture consistant à placer le portugais au premier plan de l'explication linguistique, Júlio Ribeiro présentera les transformations qui conduiront au portugais à partir du latin, afin de montrer le « principe biologique qui, avec l'action des milieux, produit la contraction » (*ibid.* : 91-92). Dans cette description, on note que, dans le développement des langues jusqu'à la dernière étape, dans laquelle se situent « toutes les langues romanes », il y a des phénomènes tels que « la persistance », « la suppression », « la chute », « la substitution », « la transformation », « l'oblitération », « l'apparition », « le passage » et « la construction » d'éléments tels qu'accents, voyelles, lettres, sons, genres, articles, cas, conjugaisons

et constructions de phrases. Dans ce processus, « l'ignorance du peuple, ou plutôt son bon sens », agit dans le sens de « réduire » les formes du langage à celles qui « exprimaient les relations de pensée les plus urgentes » et, parce qu'on ne pouvait pas « s'adapter à des formes diverses et en apparence irrégulières », le peuple adopte la forme « plus régulière, la plus homologue, la moins complexe de toutes », opération historiquement réalisée à partir de la « langue latine populaire » (*ibid.* : 95).

En mettant en parallèle l'évolution de la langue et de l'humanité elle-même à travers l'histoire, Júlio Ribeiro soutient que certains peuples vont plus loin que d'autres dans le développement des langues au cours des siècles, établissant ainsi des hiérarchies entre les langues et les locuteurs. Ainsi, à partir de la traduction d'une partie du livre d'André Lefèvre, *Études de linguistique et philologie* (1877), parlant des « progrès » et des « déviations » du développement de l'intelligence, Júlio Ribeiro raconte qu'à « surprendre dans ses rudiments extrêmes ce langage qui d'un animal a fait l'homme » on serait « en présence des cinq cents racines à peu près irréductibles », « monosyllabes », qui « amalgamées, agglutinées, fondues ensemble, greffant mot sur mot, idée sur idée, ont fourni des millions d'expressions aux millions de pensées et de nuances intellectuelles », soutenant « la trame entière de nos langues indo-européennes ». Les Chinois, une fois de plus, sont présentés comme un exemple de peuple qui « en sont restés là et qui ont suppléé par des prodiges de nuances vocales et de rigueur logique à l'indigence du système monosyllabique », lorsque l'homme « à mesure que sa langue se déliait, attachait ses rudes monosyllabes aux divers actes de sa vie grossière ». Dans cette ligne de développement du langage, on dit que « les Mongols, Finnois, Turcs, n'ont pas dépassé l'agglutination, procédé ingénieux, mais trop simple encore, qui semble caractériser aussi tous les dialectes américains et africains », tandis que « les Sémites et les Aryens ont poussé plus loin dans la même voie ; ils sont arrivés à la fusion des racines verbales et des suffixes pronominaux, à la *flexion*. » Il a fallu des siècles et des siècles, selon le récit d'unification et de dispersion linguistique de Júlio Ribeiro, « avant que l'homme articulât ses cris et nuançât son articulation » et « avant que des milliers de langues

individuelles ou familiales se soient fondues en un seul idiome, père lui-même de dialectes innombrables » (*ibid.* : 115-117).

#### 4. Évolution : mécanismes et principes du développement et progrès des langues et des peuples

C'est l'évolution naturelle qui sous-tend les mécanismes et principes du développement des langues et des peuples qui les parlent, question qui bénéficiera d'un chapitre entier du livre de Júlio Ribeiro, construit à partir de la traduction de la deuxième partie de l'œuvre *Le Darwinisme* publiée par Émile Ferrière en 1872.

En référence au géologue Charles Lyell, auteur du texte *The Geological Evidences of the Antiquity Man*, de 1868, Júlio Ribeiro déclare que « comme les espèces organiques qui peuplent le monde, les langues, véritables organismes sociologiques, sont soumises à la grande loi de la lutte pour la vie, la loi de la sélection ». Dans cette théorisation de l'évolution linguistique, la langue est affectée par le peuple : « la chance des langues est liée à la chance des peuples » (*ibid.* : 46). Il y a donc un certain nombre d'éléments de la vie humaine marqués comme « cause de variation et de sélection » des langues.

« Les relations commerciales, industrielles, politiques et littéraires » et les « progrès réalisés par les arts, l'industrie et la science » sont mentionnés pour affirmer que « les causes de sélection les plus puissantes sont politiques ou littéraires » (*ibid.* : 44-45). Cette explication ne semble pas suivre entièrement le programme du comparatisme classique, notamment le modèle des lois phonétiques. Comme l'explique Auroux (2000a), les deux premières générations de comparatistes ont fait appel à une *explication historique* pour représenter la nature et le développement des entités linguistiques à partir de métaphores organicistes qui considéraient la langue comme un organisme vivant né et grandissant selon des *lois fixes*, selon des *lois historiques*. À partir de 1870, selon Auroux (2000a), les comparatistes ont commencé à emprunter au modèle de la physique des lois pour traiter de la nature et du développement des langues, formulant ainsi la théorie des *lois phonétiques*, dans un mouvement qui cherche à placer la linguistique au sein des sciences naturelles. Appelés

néogrammairiens, les comparatistes qui travaillaient principalement avec les lois phonétiques, comme l'explique Oesterreicher (2000), exagèrent l'importance du facteur phonétique au détriment des éléments sémantiques, psychologiques et socioculturels. Tel que défini par Engler (2000), le théorème des lois phonétiques des néogrammairiens considère que les sons d'une langue à une position donnée et à une certaine période changent régulièrement et sans exception. Pour les néogrammairiens, comme le dit Auroux (2000b : 390), « une langue donnée est une espèce quasi-naturelle qui évolue automatiquement sous l'action de lois phonétiques nécessaires ». Affectant la question centrale du comparatisme, en particulier le modèle des néogrammairiens, la formulation de Júlio Ribeiro pour le changement linguistique apporte des principes explicatifs autres que les lois phonétiques, mettant les éléments sociaux, économiques, politiques et culturels au premier plan.

En ce qui concerne les effets sur la variation et la sélection linguistique des avancées apportées par les arts, l'industrie et les sciences, Júlio Ribeiro enseigne que l'invention de la presse, l'utilisation de la vapeur, l'usage des chemins de fer et l'application industrielle des découvertes faites par la physique et la chimie ont « enrichi le langage avec un grand nombre de mots et de métaphores » en introduisant de « nouvelles expressions » dans les « langues modernes », un processus évalué positivement car représentant un enrichissement, on affirme que les locutions sont devenues « obsolètes car elles ne correspondent plus aux exigences de la vie contemporaine » (Ribeiro 1880 : 44-45).

Les modifications de l'anglo-saxon par l'introduction du français en Grande-Bretagne, la consécration du Toscan sur les autres dialectes italiens en raison de la *Divine Comédie* de Dante et la fixation portugaise au XVI<sup>e</sup> siècle par *Os Lusíadas* sont des exemples donnés de la puissance de la politique et de la littérature pour la sélection des langues. Ainsi, on soutient que la conquête d'un pays « a pour résultat certain une altération plus ou moins profonde de la langue des vaincus » et que « le génie littéraire occupe une place distincte comme cause de sélection » (*ibid.* : 45).

La « destruction » des langues et des « pertes irréparables » sont deux facettes de la soumission des langues à la sélection. Selon Júlio

Ribeiro, « d'innombrables langues ont succombé sans que leur vie ait été recueillie et préservée, comme elle le serait aujourd'hui, grâce à la presse et à la diffusion des lumières », ayant, au vu des « causes de leur destruction » nombreuses et variées, très peu de « conservation des fossiles linguistiques », par exemple « la littérature grecque et latine », qui « aurait été irrévocablement perdue à l'exception des frères bénédictins » (*ibid.* : 46). Condamnant qu'« aucune langue morte ne peut revivre » (*ibid.* : 47) et que « la probabilité de leur durée est proportionnelle au nombre d'hommes qui les parlent » (*ibid.* : 46), l'auteur déplore également que « par la force des variations continues et une sélection toujours active il n'y a pas une seule langue vivante qui n'ait subi de pertes irréparables » (*ibid.* : 47). La description est plutôt négative, fataliste. L'exemple donné est la quasi-inintelligibilité « pour la langue vulgaire », « aujourd'hui », d'auteurs tels que Fernão Lopes et Azurára, Gil Vicente et Camões, dans une autre illustration de la position de Júlio Ribeiro posant la langue portugaise comme élément de démonstration de la théorie linguistique. La littérature semble être le seul aspect du passé des langues qui mérite d'être pleuré, et tout le reste doit être laissé de côté, surmonté, car, selon l'auteur, « aucun de ces fossiles ne sera ramené à la vie, l'humanité ne pourrait pas renoncer à son gré aux conditions de la vie actuelle, pour raviver les vieux préjugés, la foi naïve et aveugle, les grossières coutumes, l'ignorance et la barbarie primitive » (*ibid.* : 48). Le mot fossile porte un sens positif lorsqu'il est lié à des textes littéraires, mais il prend un caractère plutôt négatif lorsque, dépourvu d'un tel support, il se réfère simplement au passé des langues.

Du fait de la division du travail comme cause de sélection, les langues progressent et s'améliorent :

*Trabalhando em geral a selecção para o bem dos individuos, segue-se que as especies melhoram. Ora a regra de todo o progresso está na divisao do trabalho physiologico. Um animal está tanto mais elevado na escala dos seres, quanto cada uma de suas funcções se executa por organ proprio e distincto. Dá-se o mesmo com as linguas.*

*Ellas tambem fazem o progresso com o tempo, e continuam fatalmente a sua marcha neste sentido.*

*As linguas em seu começo constavam de poucas palavras, e essas palavras tinham sentido mais comprehensivo. Uma só servia para traduzir*

*muitas idéas diferentes: dahi a obscuridade do pensamento. Um dos mais attrahentes e instructivos estudos que ha é seguir as restricções que cada seculo impõe á excessiva comprehensão das palavras.*

*E' uma verdadeira divisão do trabalho intellectual operada paulatina-mente na linguagem.*

*Haja vista, por exemplo, ás accepções diversissimas que teve a palavra—genio—para vir a significar—o poder de crear.*

*Este exemplo, tomado de entre mil, mostra com bastante lucidez qual a marcha seguida pelas linguas. Esta eliminação successiva de sentidos diversos accumulados em uma só ex-pressão é uma das leis que dirigem a transformação das linguas, e que as fazem acercar-se a passos lentos mas continuos de seu idéial —Perspicuidade e Clareza. —As linguas por seu turno exercem sobre a mente do homem influencia salutar: ellas ajudam-na a operar sobre as idéias um trabalho analogo de divisão, e merecem assim que se lhes dê o nome de methodos analyticos. (Ribeiro 1880 : 48-49)*

‘Travaillant généralement sur la sélection pour le bien des individus, il s’ensuit que les espèces s’améliorent. Or la règle de tout progrès réside dans la division du *travail physiologique*. Un animal est d’autant plus élevé sur l’échelle des êtres que chacune de ses fonctions est exercée par son propre organe distinct. C’est la même chose avec les langues.

Elles font également des progrès au fil du temps et continuent leur marche fatalement en ce sens.

Les langues à leurs débuts étaient composées de peu de mots, et ces mots avaient un sens plus compréhensible. L’un a servi à traduire de nombreuses idées différentes : d’où l’obscurité de la pensée. L’une des études les plus intrigantes et instructives est de suivre les *restrictions* que chaque siècle impose à la compréhension excessive des mots.

C’est une véritable *division du travail* intellectuel opérée progressivement dans le langage.

Considérez, par exemple, les diverses acceptions du mot - *génie* - signifiant in fine - *le pouvoir de créer*.

Cet exemple, parmi mille autres, montre très clairement le cheminement que suivent les langues. Cette élimination successive de diverses significations accumulées en une seule expression est l’une des lois qui entraînent la transformation des langues et les conduisent par des étapes lentes mais continues à se rapprocher de leur idéal - *la perspicacité et la clarté*. Les langues exercent à leur tour une influence salutaire sur l’esprit de l’homme : elles l’aident à opérer sur des idées un travail analogue de division, et méritent donc d’être appelées *méthodes analytiques*.

Comme produit de deux facteurs inévitables, le temps et la division du travail intellectuel, comme l'indique l'adverbe « fatalement » utilisé dans la formulation, le progrès des langues viserait un idéal, à savoir la perspicacité et la clarté, atteints par l'élimination des significations accumulées, du raffinement, de la restriction des sens, surmontant « l'obscurité de la pensée », dont le contrepoint est la multiplication du nombre de mots. Dans ce jeu, il y a une action réciproque entre les langues et l'esprit de l'homme : le progrès des langues est l'œuvre d'un travail intellectuel opéré par les hommes, et il opère un travail intellectuel dans l'esprit des hommes.

Par la « constance de structure, que ce soit dans les racines, les organes lexicologiques de la plus haute importance, ou dans les flexions, les organes d'importance variable » les langues sont intégrées à une « communauté d'origine », comme, par exemple, l'italien, l'espagnol, le français, le portugais et la langue valaque », et certaines, « par variation et par sélection, se sont éloignées de la langue mère » (*ibid.* : 50-51). S'il était possible de marquer l'appartenance et l'éloignement des langues d'une communauté d'origine, pour Júlio Ribeiro, « le portugais moderne a conservé de nombreuses traces de son origine latine, la parenté est encore plus évidente dans la phase embryonnaire » (p. 53), tissant donc une séparation entre « phase embryonnaire » et « portugais moderne » et arguant qu'une communauté d'origine se révèle par deux « traces » : « 1) par des lettres rudimentaires ou atrophiées ; 2) par la structure embryonnaire, c'est-à-dire par la forme qu'avaient les mots et l'orthographe d'une langue avant d'acquérir une virilité stable par l'éclosion des chefs-d'œuvre de sa littérature » (*ibid.* : 51). Il convient de noter l'esquisse des deux moments de l'histoire des langues, un moment d'instabilité et un autre de stabilité, une phase embryonnaire et une phase moderne, et le passage de l'un à l'autre se ferait à travers le canon littéraire. C'est donc la littérature qui donne de la stabilité aux langues.

## 5. Conclusion

Soumises à la loi de l'évolution et à l'action de l'environnement et du temps, mais aussi en vertu de facteurs sociaux, économiques, politiques et culturels, toujours au travers du peuple qui les parle,

pour Júlio Ribeiro, les langues évoluent. Elles s'enrichissent ou sont obsolètes lorsqu'elles répondent ou non aux exigences de la vie contemporaine, qu'elles suivent ou non les progrès apportés par les arts, l'industrie et les sciences. Elles changent lorsque leurs locuteurs sont vaincus par d'autres peuples. Elles sont détruites ou subissent des pertes irréparables, selon le nombre d'êtres humains qui les parlent, laissant derrière eux des fossiles qui lèguent une littérature précieuse ou témoignent d'un passé d'ignorance et de barbarie. Elles progressent et s'améliorent du travail intellectuel vers l'idéal de perspicacité et de clarté, affectant et étant affectées par l'esprit humain. Elles se connectent avec une communauté d'origine et s'éloignent de leur langue-mère et peuvent conserver des traces de leurs origines. Elles se stabilisent avec le temps, principalement grâce à la littérature. Cependant, toutes les langues ne suivent pas cette voie au même rythme, chacune étant à un des quatre stades de développement, d'où leur regroupement et leur classification hiérarchique en quatre familles : monosyllabiques, agglutinantes, flexionnelles, contractées. Dans cette histoire du développement, les langues contiennent et présentent des éléments de plus ou moins de complexité, comme « la voix libre –a– », « de tous les éléments de la langue, c'est le moins complexe, qui se produit le plus facilement » (*ibid.* : 27).

## Bibliographie

### Sources primaires

Ribeiro, Júlio. 1880. *Traços geraes de linguistica*. São Paulo : Abilio A.S. Marques.

### Sources secondaires

Auroux, Sylvain. 2008. *Questão da origem das línguas. A historicidade das ciências*. Campinas : RG.

Auroux, Sylvain. 1992. *A revolução tecnológica da gramatização*. Campinas : Unicamp.

- Auroux, Sylvain. 2000a. Introduction : émergence et domination de la grammaire comparée. *Histoire des idées linguistiques : l'hégémonie du comparatisme*, dir par Sylvain Auroux. Liège/Bruxelles : Pierre Mardaga. 9-22.
- Auroux, Sylvain. 2000b. Les langues universelles. *Histoire des idées linguistiques : l'hégémonie du comparatisme*, dir par Sylvain Auroux. Liège/Bruxelles : Pierre Mardaga. 377-408.
- Auroux, Sylvain, Gilles Bernard & Jacques Boulle. 2000. Le développement du comparatisme indo-européen. *Histoire des idées linguistiques : l'hégémonie du comparatisme*, dir par Sylvain Auroux. Liège/Bruxelles : Pierre Mardaga. 155-172.
- De Aquino, José Edicarlos. 2016. *Júlio Ribeiro na história das ideias linguísticas no Brasil*. Thèse de doctorat. Universidade Estadual de Campinas & Université Sorbonne Nouvelle – Paris 3. [en ligne : <https://theses.hal.science/tel-01541881>]
- Desmet, Piet. 1996. *La linguistique naturaliste en France (1867-1922) : nature, origine et évolution du langage*. Louvain/Paris: Peeters.
- Engler, Rudolf. 2000. La géographie linguistique. *Histoire des idées linguistiques : l'hégémonie du comparatisme*, dir par Sylvain Auroux. Liège/Bruxelles : Pierre Mardaga. 239-252.
- Hovelacque, Abel. 1877. *La linguistique*. Paris : C. Reinwald et Cie.
- Lefèvre, André. 1877. *Études de linguistique et de philologie*. Paris : Ernest Leroux.
- Klippi, Carita. 2010. *La Vie du langage : la linguistique dynamique en France de 1864 à 1916*. Lyon : ENS Éditions.
- Lyell, Charles. 1868. *The Geological Evidences of the Antiquity Man*. Londres : John Murray.
- Müller, Max. 1855. *Lecture on the Science of Language*. Londres : Longmans, Green and Co.
- Oesterreicher, Wulf. 2000. L'étude des langues romanes. *Histoire des idées linguistiques : l'hégémonie du comparatisme*, dir par Sylvain Auroux. Liège/Bruxelles : Pierre Mardaga. 183-192.
- Renan, Ernest. 1858. *De l'origine du langage*, 2<sup>e</sup> éd. Paris : Michel Levy.
- Schmitter, Peter. 2000. Le savoir romantique. *Histoire des idées linguistiques : l'hégémonie du comparatisme*, dir par Sylvain Auroux. Liège/Bruxelles : Pierre Mardaga. 63-78.
- Trabant, Jürgen. 2000. Le courant humboldtien. *Histoire des idées linguistiques : l'hégémonie du comparatisme*, dir par Sylvain Auroux. Liège/Bruxelles : Pierre Mardaga. 311-322.

# La langue et la pensée



# Porphyry's theory on linguistic simplicity and complexity in the frame of Neoplatonic teaching

**Maria Chriti**

Aristotle University of Thessaloniki, School of Philology, Department of Classics

---

## ABSTRACT

In this paper an attempt is made to contextualize the views of the Neoplatonic commentator Porphyry concerning language's progression from simplicity to complexity, into the curricula of the Alexandrian School of philosophy, in terms of proceeding from simple to complex philosophical reflection and reasoning. "Simple words" are considered by Porphyry and later Neoplatonic commentators as the subject-matter of the *Categories*, while the "more complex" level of human signification, i.e., the "second imposition of words", is considered as the subject-matter of the second logical work, i.e., *On Interpretation*. Given the established sequence of the treatises in the *Organon* within the Neoplatonic curricula, the importance of this linguistic step from simplicity to complexity can be related to the respective progress in thinking and reasoning.

## KEYWORDS

Porphyry, simplicity, complexity, language, thinking, Neoplatonism

## RÉSUMÉ

Cet article discute les vues du commentateur néoplatonicien Porphyre, concernant le passage de la simplicité à la complexité linguistique, dans le cadre du programme de l'école philosophique d'Alexandrie. Les « sons vocaux simples » sont considérés par Porphyre et plus tard par les commentateurs néoplatoniciens comme le sujet des *Catégories*, tandis que

le niveau « plus complexe » de la signification humaine, c'est-à-dire la « deuxième imposition des mots » est considérée comme le sujet du deuxième ouvrage logique, *Sur l'interprétation* ; étant donné la séquence établie des traités de l'*Organon* dans les programmes néoplatoniciens, l'importance du passage de la simplicité à la complexité dans le langage peut être liée aux progrès respectifs de la pensée et du raisonnement.

#### MOTS-CLÉS

Porphyre, simplicité, complexité, langage, pensée, néoplatonisme

---

## 1. Introduction: the Neoplatonic School of Alexandria

This paper attempts to contextualize the views of the Neoplatonic philosopher and commentator Porphyry as relating to the progress from linguistic simplicity to linguistic complexity, into the curricula of the School of Alexandria.

The Neoplatonic School of Alexandria is a particular case in the history of ancient philosophical schools, having affected reflection through Late Antiquity and Byzantium, as well as the Islamic East and Latinized West in an irreversible way<sup>1</sup>, not only with its doctrines, but also with its influential representatives<sup>2</sup>. Those scholars transmitted the study of philosophy in their own way, as framed by commenting on earlier great thinkers. Therefore, delving into the stages of such a curriculum can illustrate a significant aspect of the process of philosophical reflection during a transitional era.

In general, when approaching Neoplatonic commentators, it is necessary to take into account the particular character of their texts, as contextualized within the curricula of the Neoplatonic Schools of Athens<sup>3</sup> and Alexandria<sup>4</sup>: Porphyry's *Introduction* to Aristotle was

---

1. As Sorabji (2005: 1) describes.

2. See Chriti 2022.

3. Plutarch of Athens was the founder of this School, which was closed at Justinian's order in 529 AD: see Beaucamp (2002) and Watts (2004); see also Sorabji (2005: 9).

4. Concerning the School of Ammonius, son of Hermeias, see Sorabji 1990; see also Westeringet *al.* (2003: x-xlii); Blumenthal (1993: 307-325). On Ammonius

the first text to be explained in these Schools. With few exceptions, Neoplatonic commentaries are structured as notes on courses from Porphyry onwards<sup>5</sup>, a turning point depicted by Sten Ebbesen<sup>6</sup>. After Porphyry's *Introduction*, Aristotle's treatises were taught as the *minor mysteries* (ἐλάσσονα μυστήρια) starting from the *Categories*, and the students would then proceed to the Platonic dialogues as the *major mysteries* (μείζονα μυστήρια) – Aristotle was “Plato's most faithful student” and he had to precede Plato, so as to contribute to the better understanding of his teacher's philosophy, as Plato was the Neoplatonists' main interest<sup>7</sup>.

Of high interest for this present study is the approach to the scale via which a philosopher-to-be was supposed to acquire the necessary qualifications, having the *Categories* at the very beginning of this specific syllabus. Aristotle's first logical treatise has caused serious debates from antiquity to contemporary scholarship<sup>8</sup> regarding its subject-matter: What is the philosopher talking about in the *Categories*? The actual purpose of the *Categories* is beyond the scope of this paper, with the focus here being the manner in which Porphyry's theory concerning this specific purpose is integrated into the educational *milieu* of the Alexandrian School: that is in terms of proceeding from lower to higher levels of thinking and signifying. This represents an innovative approach to Porphyry's legacy to logic and semantics, and while the excellent contributions of Sten Ebbesen (1990) and Voula Kotzia (1992: 23 ff) primarily investigate

---

see Saffrey (1989: 168-169); Verrycken (1990); Blank (1996: 1 ff). The School of Athens mostly emphasized Plato, while the School of Alexandria emphasized Aristotle; in general, however, the two Schools did not have crucial differences as regards their method and orientation: see Golitsis (2008: 9, note 9).

5. On the issues of the Neoplatonic teaching in the respective Schools see Sorabji (2005: 1-28).

6. Ebbesen (1981: 133). Ebbesen's view is also adopted by Kotzia (1992: 21).

7. For a concise overview of the specific series of courses designed to prepare the Neoplatonic philosopher, see Golitsis (2008: 10-16), with further bibliography. On the two fundamental principles of approaching Aristotle and Plato in these schools see in general Karamanolis (2006); see also Sorabji (1990: 3) and Kotzia (2007: 194-201).

8. See Chriti (2019a: 419).

his *Introduction* and surviving commentary on the *Categories*, the former aspect has not been discussed.

Porphry's lost and more extended commentary *To Gedaleius* can be reconstructed thanks to the evidence provided by Simplicius, a text not written according to any oral teaching but rather to be read (πρὸς ἀνάγνωσιν)<sup>9</sup>. Simplicius' teacher was Ammonius of Hermeias (5<sup>th</sup>-6<sup>th</sup> cent. AD), Head of the School, according to the teaching of whom a number of later commentaries on the *Categories* were written. Thus, John Philoponus' (6<sup>th</sup> cent. AD) commentary reflects Ammonius' teaching<sup>10</sup>. After Philoponus, we will examine what Olympiodorus (6<sup>th</sup> cent. AD) says in his respective commentary<sup>11</sup>. Another commentary on the *Categories*, this time under David's name (CAG XIII), was believed to have been written by Elias, a Christian student of Olympiodorus<sup>12</sup>; the identity of Elias as the author of this text is followed here, and is examined last. The manner in which Porphyry's theory integrates into the syllabus of Alexandria's Neoplatonic School, as an aspect of how to approach the correlations between thinking and language for a philosopher-to-be of that era, will be illustrated in the Conclusions.

## 2. Porphyry on simple words in the first human linguistic act

The student, editor and biographer of Plotinus, argued that the subject of the *Categories* is “the first imposition of words”<sup>13</sup> (“πρώτη θέσις

---

9. See Praechter (1922: 507-508); Westerink (1962: xxvii). Concerning the way Simplicius exploits his “rich philosophical bibliography”, see Chase (2003: 1).

10. See the analysis by Kotzia (1992: 139 ff).

11. The commentary on the *Categories* under Olympiodorus' name is based on his teaching (ἀπὸ φωνῆς Ὀλυμπιοδώρου) and contains linguistic views which certify linguistic expositions in other commentaries: see Kotzia (1992:183).

12. Hadot (1987: 5, note 7) supported the attribution to David. In any case, David and Elias were both Christian commentators (second half of the 6<sup>th</sup> cent. AD), having a similar philosophical orientation. The basic indication for Elias' studentship is the high number of parallel texts between his commentary and that of Olympiodorus.

13. Porph. *On Cat.* 58.5 & 58.32-33.

τῶν ὀνομάτων”)<sup>14</sup>, that is, the first words as established for things by mankind. As Porphyry describes, things pre-existed human beings, who were faced with the need to declare (κατονομάζειν) and signify (δηλοῦν) them “also with their voice” (καὶ διὰ τῆς φωνῆς; man is a δηλωτικὸν ὄν, he says, a “declarative being”)<sup>15</sup>. Porphyry’s stressing “also with his voice” is clarified when he uses the deictic phrase “τόδε τί”, so as to explain that people named certain things which were before their eyes and they could show<sup>16</sup>. In mankind’s first linguistic act, there is nothing divine or supernatural: rather they used their articulated vocal sounds and decided to call something a “dog”, something else a “pedestal”, something else a “man”, another thing “the sun”, one colour “white”, another colour “black”, something else as a “number”, another as a “size”, etc. (Porph., *On Cat.*,57.23-29):

---

14. Ὄνομα here has the sense of ‘word’, as it is ‘said in two ways’ (διχῶς λεγόμενον) in ancient linguistic thinking, according to Aristotle: *a*) it is any meaningful utterance (φωνὴ σημαντική); *b*) in a *categorical statement*, a *name* designates the agent of a verb to the action of the verb, the *subject* (*On Interpretation*, 16b19-20).

15. Porph., *On Cat.*,57.20-23. The outlines of Ebbesen (1990: 382) and Kotzia (1992: 21-31) are in general followed in this article to approach Porphyry’s text. What Ebbesen has marked out regarding the not strict use of linguistic terminology in philosophical texts of this period can be certified regarding the terms κατονομάζειν and δηλοῦν, which stand for human linguistic activity in Porphyry’s theory: the first term is only used by Porphyry in this text, obviously with the sense ‘to name’ (1990, p. 156-157), while the form δηλοῦν is used by him several times in his texts, in most of which it has the sense ‘to signify’ for words and linguistic expressions; see for example *On Cat.*, 62.8, 107.20; *On Tim.*, 1.21, 14.9; *Hom. Quest.*, 2.447.17, 8.1.46 etc. In the rest of the occurrences the term is used to designate ‘represent/signify’ for signs and actions, but in no case does Porphyry use δηλοῦν for simple sounds (ψόφοι), or even for inarticulate vocal sounds. If we also take into consideration that the term φωναί often means ‘words’ in his commentary on the *Categories* (see for example *On Cat.*, 56.35, 58.13, 58.30, 62.31, 71.20, 96.11, 102.8) we can conclude that in this text Porphyry is not dealing with the distinction between language and inarticulate vocal sounds or mere sounds (ψόφοι), but the distinction between non-vocal designation of things on the one hand, and vocal designation of things via articulated human sounds on the other..

16. Porphyry says that the word, e.g., *stone*, can be attributed to the very specific stone that we perceive by means of our senses: “τοῦδε τοῦ δεικνυμένου λίθου” (Porph., *On Cat.*, 56.7-13).

καὶ γέγονεν αὐτῷ ἡ πρώτη χρῆσις τῶν λέξεων εἰς τὸ παραστήσαι ἕκαστον τῶν πραγμάτων διὰ φωνῶν τινῶν καὶ λέξεων, καθ' ἣν δὴ σχέσιν τῶν φωνῶν τὴν πρὸς τὰ πράγματα τότε μὲν τι πρᾶγμα βάθρον κέκληκεν, τότε δὲ ἄνθρωπον, τότε δὲ κύνα, ἥλιον δὲ τότε, καὶ πάλιν τότε μὲν τὸ χρῶμα λευκόν, τότε δὲ μέλαν, καὶ τότε μὲν ἀριθμόν, τότε δὲ μέγεθος, καὶ τότε μὲν δίπηχυ, ... καὶ οὕτως ἕκαστω πράγματι λέξεις καὶ ὀνόματα τέθεικεν σημαντικὰ αὐτῶν καὶ μηνυτικὰ διὰ τῶν τοιούτων τῆς φωνῆς ψόφων.

Thus his first use of words (ἡ πρώτη χρῆσις τῶν λέξεων)<sup>17</sup> came to be to represent each thing by means of certain vocal sounds and words. In accordance with this relation between words and things, this thing here (τότε μὲν τι) is called a 'chair', that a 'man', this a 'dog', that 'the sun', and again, this colour is called 'white', that 'black', and this is called 'number', that 'size', this 'two cubits', and that 'three cubits'... In this way words and names have been assigned to each thing, serving to signify and communicate them by employing such sounds of the voice. (adapted transl. of S. K. Strange)

It can be said that Porphyry suggests that the most primitive level of declaration is the *deictic* one, during which no vocal sound is produced and human beings gesture to things that are in front of them. Be that as it may, the next level concerns words which represent experiential data. As this specific stage is cited immediately after the action of "pointing to", it can be deduced that the Neoplatonic philosopher refers, not only to the act of "talking about things which are present", but also to cases of people referring to things which are not present. This is demonstrated by his use of the term "συμβολικῶς"<sup>18</sup>, which is employed by Aristotle in his *Sophistical Refutations*<sup>19</sup> to designate precisely this substitution of things and generally the conventional

---

17. The translation of S. K. Strange is not followed here; Porphyry probably uses "φωναί" and "λέξεις" alternatively to denote "words", although Strange translates "λέξις" as 'expression'; it wouldn't be unreasonable to suppose that it means 'word' in the Modern Greek sense (λέξη), since Porphyry seems to use the term in the specific sense as many times as with the sense 'linguistic expression' in his texts: see for example *On Cat.*, 55.12, 56.7, 58.8, 59.11, 61.27, 75.3, 86.21, 86.37, 91.8-19; *Hom. Quest.*, 4.2.6, 12.10.53, 19.221.7 etc.

18. He uses the term a little further down, in 57.30.

19. *Sophistical Refutations*, 165a6 ff.

character of words that represent things in *On Interpretation*<sup>20</sup>. Therefore, the fact that human beings imposed words that could be used as *tokens* of the things not present, demonstrates a very significant step of the human mind, i.e., from reflexive to symbolic representation, something which is pre-supposed for the following level of declaration.

Before embarking on this final level, it should be highlighted that Porphyry notes that not all words are examined in the *Categories*, but only 'simple words which signify things', i.e., words "to the extent that" (καθό) they represent something. Simplicity is the essential feature of the first linguistic representation of things, as it is a non-complex act to represent an object. Words were imposed as tokens of things, in accordance to human conceptualization, which means that these words do not reflect the things' existence itself, but rather the way in which human mind conceives them<sup>21</sup>, since they exist independently of human perception: indeed, things are there, whether human beings can perceive them or not.

The most important parameter in this aspect is that mental activity is included in the first imposition of words, even if it is not openly stated by Porphyry, as Kotzia (1992: 47 ff) wisely argued. It is possible that Porphyry explicitly included mental entities in his lost commentary on the *Categories* when investigating the treatise's purpose, as well as in his lost commentary on Aristotle's *On Interpretation*<sup>22</sup>. Nevertheless, all three parameters, i.e., things, mental entities and words, are present in the formulations of later commentators to express the subject matter of the first logical treatise, possibly also under the influence of Aristotle's "semantic triangle" in *On Interpretation*<sup>23</sup>. Symbolic signification and signification through a concept, are not in opposition with each other, as the existence of three parameters of signification (words, concepts, things) may

---

20. On this specific discussion see Chriti (2018).

21. For this particular distinction see Porph., *Intr.*, 1.10-11 & *On Cat.*, 75.27-28, as well as the analysis by Kotzia (1992: 45).

22. See Ebbesen (1981: 151-152).

23. Amm., *On Cat.*, 9.17-18; 9.22; 10.3; 10.8; 10.13; 12.1. Simpl., *On Cat.*, 12.1 ff; Philop., *On Cat.*, 10.6-8; Olymp., *On Cat.*, 69.15-17; Elias, *On Cat.*, 170.15-18 & 176.33-34. For this specific discussion see Chriti (2021: 44-45).

already be detected in Porphyry who uses the term *συμβολικῶς*. The existence of a notion/concept does not exclude the concept of ‘symbolic/ conventional representation’.

However disputable Porphyry’s approach to the subject of the *Categories* may be considered today, he discusses the primitive stages of mental elaboration and linguistic representation respectively, by describing what happened when human beings realized that they should stand among the data of reality and communicate them. The first human linguistic act results in words which are related to things, since human beings use their voice as a tool to symbolize objects of their perception and this is how human development in thinking and language begins.

### 3. Composition in the second human linguistic act

The conventional and communicative character of words in the “first imposition” provided human beings with another privilege in terms of linguistic use, elevating mankind to a superior level of representation. After the initial declaration of things, speakers reconsidered the function of words in speech and evolved into another mode of use as related to potential combination. Indeed, the “second imposition of words” (*δευτέρα θέσις τῶν ὀνομάτων*) is a categorization no longer related to the simple representation of things, but to discussion of language itself.

The revisiting of the same words — but this time in combination — is described as follows by Porphyry: words preceded by an “article” were called *ὀνόματα* (‘names’), while those which could be inflected in certain ways were called *ῥήματα* (‘verbs’). The secondary approach to words constituted mankind’s next act of declaration and this second “name-giving” resulted to the utterances by which mankind could now refer to other utterances and, to be more precise, to the concepts of other utterances; this was a “language about language” (‘metalanguage’; Porph., *On Cat.*, 57.29-58.4)<sup>24</sup>:

---

24. See also Amm., *On Cat.*, 11.15-12.1: *κατὰ δευτέραν δὲ ἐπιβολὴν ἐπεσκέψαντο ὅτι ταῖς μὲν τῶν φωνῶν δύναται συντάττεσθαι ἄρθρα, χρόνοι δὲ οὐχί, ἅπερ ἐκάλεσαν ὀνόματα, ταῖς δὲ χρόνοι μὲν συντάττονται, ἄρθρα δὲ οὐ, ἅπερ εἰσι*

τεθεισῶν δὲ τοῖς πράγμασι συμβολικῶς τινων λέξεων προηγουμένως, πάλιν ὁ ἄνθρωπος κατὰ δευτέραν ἐπιβολὴν ἐπανελθὼν αὐτὰς τὰς τεθείσας λέξεις θεωρήσας τὰς μὲν τοιοῦτον φέρε τύπον ἔχούσας, ὥστε ἄρθροις συνάπτεσθαι τοιοῖσδε, ὀνόματα κέκληκε, τὰς δὲ τοιαύτας οἶον τὸ περιπατῶ, περιπατεῖς, περιπατεῖ, ῥήματα, δηλώματα τῶν ποιῶν τύπων παριστὰς τῶν φωνῶν διὰ τοῦ τὰς μὲν ὀνόματα καλέσαι τὰς δὲ ῥήματα. ὥστε τόδε μὲν τι τὸ πρᾶγμα καλέσαι χρυσὸν καὶ τὴν τοιαύτην ὕλην τὴν οὕτω διαλάμπουσαν προσαγορευῆσαι ἥλιον τῆς πρώτης ἢν θέσεως τῶν ὀνομάτων, τὸδὲ τὴν χρυσὸν λέξιν εἰπεῖν εἶναι ὄνομα τῆς δευτέρας θέσεως καὶ τοὺς τύπους τῆς ποιᾶς λέξεως σημαινούσης.

When certain words had been credited to things in the first place, man began to reflect upon the words that had been posited from another point of view, and saw that some were of such a form as to be attached to certain articles: these he called 'nouns'. Others, such as 'walk' and 'walks', he called 'verbs', indicating the qualitative differences between the two types of words by calling the one 'nouns' and the other 'verbs'. Thus, calling this sort of thing 'gold' and that material that shines so brightly 'the sun' belongs to the primary imposition of words, while saying that the expression 'gold' is a noun belongs to their secondary imposition, which signifies the qualitatively different types of expression. (adapted transl. of S. K. Strange)

Porphyry formulates a rather complete theory concerning a rank of certain levels of human thinking and declaration, the profundity of which is crucial for the conceptualization of the process from simplicity to complexity in mental and linguistic activity. The first level was deictic which preceded even the primary linguistic declaration of a thing. The next was naming things with simple words (the first linguistic act). The last level was the approach to the combination of words as the superior linguistic act, a procedure which is distanced from simplicity. This final level of declaration is to "use words to talk about words", a stage that concerns language by which an abstract discussion is possible and now mankind can discuss concepts and not simply sensible things:

---

ῥήματα. Philoponus notes that the action of *imposing a name* on each thing did not automatically lead to any kind of distinction between *names* and *verbs*, which was a further step (Philop., *On Cat.*, 11.34-12.3).

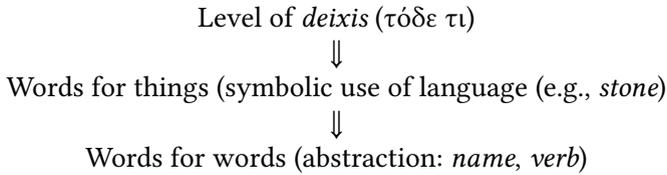


Fig. 1: Porphyry's "scale of declaration"

Porphyry started to evince his attitude towards human "declaration-levels" from implying the non-presence of language and ending with the potential for an abstract use of language. This is done via the description of a mental evolution, since human beings passed from showing objects to thinking and talking about objects without them being present and then to theorizing the use of those words in their own right, and classifying utterances as either *names* or *verbs*. Porphyry conceives of the ability to use words to refer to other linguistic units, or rather "to refer to concepts", as an inherent benefit to mankind. This is evident in his citing two examples that depict these powers<sup>25</sup>: the Greek word χρυσός ('gold') as a *signifier* of something specific (τόδε τι), i.e., the metal, belongs to the "πρώτη θέσις"; the same word as a *name* belongs to the "δευτέρα θέσις". The same is true of the word ἥλιος signifying the 'sun' that shines as a word of the "first imposition". With these two examples, Porphyry delineates the importance of his theory concerning the evolution to linguistic abstraction: what he apparently means is that it is one thing to say "my ring is *gold*" ("first imposition") and another thing to say "*gold* is a name" ("second imposition")<sup>26</sup>: in the second case we refer to a concept by talking theoretically and not in relation to a specific object. Again, the first mental and linguistic act is a simple one, while the second is complex because human beings are distanced from the world of senses and they can actually be engaged in a theoretical conversation.

As already said, according to Porphyry, the "first imposition of words", i.e., words that symbolize things, constitute the subject-

25. See the ancient text right above and especially Porph., *On Cat.*, 58.4-8.

26. Regarding the non grammatical character of the "second imposition of words" see right below.

matter of the *Categories*. The *Categories* deal with simple words “to the extent to which they *signify* things” and not from the perspective of any kind of function within human speech: “καθό σημαντικαί εἰσι τῶν πραγμάτων”. From the vantage point of Porphyry, in his *Categories* Aristotle refers to a primary categorization of experiential data with the help of language, which provides the philosopher with terms for each category<sup>27</sup>. As for the “second imposition”, this is the subject matter of the second Aristotelian logical treatise, i.e., *On Interpretation* (Porph., *On Cat.*, 58.34-36):

In *On Interpretation* the discussion concerns the 2<sup>nd</sup> imposition, which does not regard words as representing things... (own translation)

In *On Interpretation* Aristotle deals with the basic parts of categorical speech (ἀποφαντικός λόγος), i.e., subject and predicate, affirmation, negation, opposition, contradiction and, in other words, what we would today call *syntax*, meaning that Porphyry considers this specific treatment as immediately following the declaration of things, treated in the *Categories*.

Obviously, *names* and *verbs* belong to Porphyry's *metalanguage* but it is important to emphasize here that this is not a “grammatical theory”: *names* can include both *nouns* and *adjectives* like *sun* and *gold* while *names* and *verbs* are classified together, based on the fact that they render human beings capable of referring to language<sup>28</sup>. Porphyry and the other Neoplatonic commentators, who faithfully follow his theory, are well aware of the grammatical doctrines according to which *names* and *verbs* are treated in the respective grammatical categories. The Neoplatonists often refer to the Stoics and the Grammarians, and in a great many cases they discuss *names* and *verbs* as “parts of speech” in their commentaries<sup>29</sup>. In Porphyry's consideration, *names* and *verbs* of the “second imposition” correspond to the *subject* and the *predicate* in *categorical speech*, as they are considered by Aristotle in *On Interpretation*, the work which

---

27. See the discussion in Lallot & Ildefonse (2002: 328-330).

28. The classification only of names and verbs in the “second imposition of words” is a long discussion which goes beyond the scope of this article. For such a discussion see in general Chriti 2019b.

29. See Chriti (2019b: 22-24).

examines a more complex process of the human mind, in comparison to the *Categories*. According to Porphyry's classification, the subject matter of the *Categories* is less complex than that of *On Interpretation* and, to be more precise, through his lens, the *Categories* is pre-supposed for *On Interpretation*.

#### 4. Porphyry's theory as integrated in the Neoplatonic School of Alexandria

Porphyry's theory on the "first" and the "second imposition of names" is adopted by the Neoplatonic commentators as a hermeneutic tool to elucidate the "purpose" of both the first two logical treatises. Let us start with Ammonius of Hermeias, who was the Head of the School of Alexandria and teacher of later Neoplatonists. In his commentary on Aristotle's *Categories*, he justifies the sequence of the philosopher's logical treatises, insisting on the direct transmission from the first to the second treatise by stressing the exclusivity of this specific sequence, on the basis of Porphyry's theory (Amm., *On Cat.* 14.20-15.1)<sup>30</sup>:

And those are opposed to the clarity and the order, since Aristotle taught that the first imposition of simple words has to precede the treatment of names and verbs; in the *Categories* he teaches us about the position of simple words, while in *On Interpretation* about names and verbs in a way that each treatise is intrinsically related with the other: because neither is it possible to study anything else before *On Interpretation* than the *Categories*, nor after the *Categories* any other text than *On Interpretation*. And the same reasonable order exists up to the *Analytics*, the treatise of argument. (own transl.)

Ammonius states that anyone who wants to investigate arguments has to proceed from simple words to the parts of a sentence and any scholar who wishes to delve into the latter will inevitably start from the simple words of the *Categories*. As a teacher, Ammonius' depiction of the progressive steps from simple to complex mental and linguistic activity is entrenched in the educational *milieu* of his

---

30. English translations without the ancient Greek texts are cited from now onwards, due to space limits.

School, with anyone instructed in logic having to reach the level of syllogisms, the subject matter of the *Analytics*.

Ammonius (as well as the other commentators) believes that, as a whole, the concept of the *Organon* is a complete instrument by which access to philosophical thinking is provided, a tool via which students could reach the utmost logical construction, namely the syllogism.

After Ammonius, and in his commentary on the same treatise, Simplicius attests that Porphyry also refers to Boethus<sup>31</sup>, stating that the *Categories* have a more simple subject matter than *On Interpretation* (Simpl., *On Cat.*, 11.23-29):

Porphyry also adds the remarks of Boethus, which are full of sharp-wittedness and tend in the same direction as what has been said. He too says that with regard to nouns and verbs, the division takes place as far as the elements of speech, but according to the categories the division takes place in so far as words are related to beings, since they are significant of the latter. (adapted transl. by M. Chase)

Therefore, Simplicius who has Porphyry's lost commentary in his hands provides us with the evidence that Porphyry may not have been the first to connect the two logical treatises via their subject matters, but that he may convey an earlier tradition from the direct circle of the Peripatos School. A little below this, Simplicius explains the reasons why Aristotle's logical treatises should be classified the way they are, for purposes of correctness, as regards the steps to logic (Simpl., *On Cat.*, 15.13-21):<sup>32</sup>

So it is natural that we start from the *Categories*, because we are thus introduced to signifying speech and to things that are signified, from the more simple to the more complicated, since after simple words we learn what is a name and a verb, then what is affirmation and negation and which are their differences; these are what we are taught in *On Interpretation* and then we come to know what is a term, a premise and an argument and which are the kinds of argumentation and the forms

---

31. Boethus was a student of Andronicus of Rhodes and Head of the Peripatos School (1<sup>st</sup> cent. BC). On Boethus see Barnes, Bobzien, Flannery & Ierodiakonou (1991: 6).

32. Let us just point out here that the "elements of speech" should not be identified with the later grammatical "parts of speech": see Chriti (2019b: 22 ff).

of each one and the methods of instruction according to each form in the *Prior Analytics*. It is in this way that we are introduced to the treatise of proof, the one which Aristotle named as *Posterior Analytics*... (own transl.)

A question naturally arising from the first line of the text above is “who is *we*?”, as Simplicius is talking in first plural person. The most probable answer would be that he is collectively referring to philosophers like himself and, more generally, to those interested in approaching philosophy, given the educational environment of the School where he is active. Simplicius defends the order that a philosopher-to-be should follow in his studies just like Ammonius and we have the impression that he is referring to the way he was initiated to philosophy himself. He does not fail to acknowledge the importance of progressing from simplicity to complexity for anyone who wants to access logic, by reaching the treatment of arguments in Aristotle’s *Analytics*.

The same mentality is also expressed by John Philoponus in his commentary on the *Categories*, where he explicitly talks about the “instrument of logic” (Philop., *On Cat.*, 10.17-11.30):

The philosophers came in need of an instrument that would help them distinguish truth from falsity and good from bad; and what is this? The syllogism ... it is impossible to talk about it before explaining what argumentation is in general, for the discussion of which we need to know what is a premise ... but we cannot understand what is a premise, i.e., a sentence, without having previously explained names and verbs, from which every speech is composed. Therefore, he firstly discusses simple words in the *Categories*, then he treats names and verbs in *On Interpretation*, afterwards he deals with argumentation in general in his *Prior Analytics* and then he focuses on syllogism in the *Posterior Analytics*. (own transl.).

John Philoponus provides the same formulaic explanation for the subject matters of the logical treatises but also for the significance of starting from simple words and processing to complex combinations, in order to acquire the necessary knowledge of logic.

Continuing with Olympiodorus, it is evident that he follows the same line as his predecessors in his commentary on the same treatise (Olymp., *On Cat.*, 21.18-25):

Well, in order to distinguish these two, it should be added that in the *Categories* we refer to the 'first imposition', while in *On Interpretation* these words are approached to as belonging to the 'second imposition'. (own transl.)

Olympiodorus was succeeded by his student Elias who adopted Porphyry's terminology and conceptualization in terms of proceeding from simple to complex in relation to that a balanced access to logic can be achieved when studying philosophy (Elias, *On Cat.*, 131.18-132.21):

...and the purpose of the *Categories* concerns simple words signifying things via concepts. ... here Aristotle does not investigate names and verbs, but he does so in *On Interpretation*. In the *Categories* the purpose regards the first imposition of simple words which signify simple things through simple concepts ... i.e., the treatise deals with the first and not the second imposition of words, which includes names and verbs. ... And if the first imposition does not exist, there can be no second imposition, while without the second imposition there is no sentence, without which there is no argument. ... Consequently, the *Categories* should be before any other logical treatise, just like simple words pre-exist anything else. (own transl.)

Elias continues the tradition during late 6<sup>th</sup> cent. AD, a period in which the teaching of the Neoplatonic School of Alexandria still survived, in contrast with the decline and closure of the other Neoplatonic schools, owing to various historical (political, social, religious) circumstances in an era marked by the emergence of ever more complex cultural, religious, political, and philosophical networks. Porphyry's influence in terms of the smooth process from simplicity to complexity for anyone wishing to access logical thinking is still persist in the writings of the first Christian teacher of the School<sup>33</sup>.

---

33. A detailed investigation of Elias' commentary regarding the survival of such doctrines would exceed the scope of this paper and could actually constitute the subject of an independent survey: it is hoped that such an appreciative study is still to come.

## 5. Concluding remarks

Let us summarize: Ammonius of Hermeias, Simplicius, John Philoponus, Olympiodorus and Elias, who represent the Neoplatonic School of Alexandria between the 5<sup>th</sup>-6<sup>th</sup> cent. AD, adopt Porphyry's theory in terms of the subject matters of the first two Aristotelian logical treatises. They likewise justify the sequence of the rest of the logical works, according to the methodology of starting from simplicity and progressively evolving to a complexity in thinking and language. Although Porphyry never directly refers to the necessity of such a process for a student, these other commentators explicitly formulate the integration of Porphyry's theory into the curriculum of their School. Not only is the title of the "Organon" fully adopted by those philosophers, but Porphyry's description of mankind's linguistic evolution is stated to be identified with the process itself from sensory perception to philosophical thinking. This prescriptive sequence reveals the intellectual preparation that was considered as being proper and necessary for the philosophers-to-be.

Therefore, the *Categories* has "simple words which signify things via concepts" as its subject matter, while *On Interpretation* focuses on *names, verbs* and *sentences* (speech = λόγος) which can be rendered as *true* or *false*. The "second imposition of words" is, according to Neoplatonic theory, the subject matter of Aristotle's *On Interpretation*, because the emphasis is on the potential combination and complexity of words. The commentators argue that this is why the given order of treatises in the *Organon* is correct the way it is, because the natural sequence for anyone who wishes to be introduced to logic is to first examine simple words and then those that construct more complicated units.

In his commentary on Aristotle's *Categories*, Porphyry formulates a theory on the way that the first words were attributed to things by human beings. Porphyry's "first imposition of words" regards simple utterances invented for things surrounding mankind, things that they could actually point to. However, after the "deictic" level of declaration, mankind reconsidered words in regard to their form and potential combination: the "second imposition of names" is a categorization of words according to this potential, which depends

on their form that renders them apt for combinations within speech. The revisiting of the same words was now based on lack of simplicity, as Porphyry explains. This second “name-giving” resulted to the language by which mankind could now refer both to concepts and language itself, explaining the process and progress from a “core language” to combinations and complexity that could lead the way to philosophy, the ultimate mental activity. Simple words can only be used to represent the data of experience, while linguistic complexity depicts the higher mental level which is demanded to reach philosophical thinking. Human progress in language is thus linked to a more sophisticated way of thinking, which goes beyond the simple deictic pointing of things.

Porphyry's theory fits perfectly with the Neoplatonic concept of the process through which a philosopher-to-be could reach Plato's theology, since the student should begin from Aristotle's *Organon* of logic with the *Categories* as the first step: simple utterances should then be followed by the parts of a sentence and then the young scholar should continue with syllogisms, which presuppose the expertise of sentences. In Plato's Academy, the itinerary would start from Mathematics and the goal concerned the Ideas; in the Neoplatonic School they should start from Plato's most emblematic student and his doctrines of logic, before embarking on the divine Platonic philosophy. To emphasize logic, what could be more normal than starting from simple utterances before delving into the formulation of syllogisms? The reasonable sequence of logical treatises is justified through the story of the invention and imposition of words, a story fully integrated in the educational environment of the School of Late Antiquity. The commentators take the classification of Aristotle's works by Andronicus of Rhodes for granted, with none of them questioning the “firstness” of the *Categories* in Aristotle's “Organon”. The appeal of Porphyry's theory is evident: Ammonius, Simplicius, Philoponus, Olympiodorus and Elias all directly or indirectly adopt Porphyry's theory that:

- Simplicity in language is identified with simplicity in thinking and the first step in accessing philosophy.
- The more complex linguistic declaration becomes, the closer the human mind is to abstraction and philosophical reflection.

- Linguistic development can represent nothing but intellectual evolution, as shown by the process and progress from naming things in front of us to deduction and demonstration.

Regardless of the fact that the sequence of Aristotle's treatises was not conceived of by the philosopher himself and no matter the correctness of the prescriptive purposes for which this sequence was applied, it is possible that we are dealing with the first explicit theorization of the intrinsic relation between linguistic and philosophical evolution from simplicity to complexity in the history of the philosophy of language.

## References

- Barnes, Jonathan, Susanne Bobzien, Kevin S. J. Flannery & Katerina Ierodiakonou, eds. 1991. *Alexander of Aphrodisias. On Aristotle Prior Analytics* 1.1-7. London: Duckworth.
- Beaucamp, Joëlle. 2002. Le Philosophe et le joueur. La date de la « fermeture de l'École d'Athènes ». *Mélanges Gulbert Dargon*. Paris: De Boccard. (Travaux et Mémoires, 14). 21-35.
- Blank, David. 1996. *Ammonius on Aristotle's "On Interpretation"* 1-8. Cornell: Duckworth.
- Blumenthal, Henry Jacob. 1993. Alexandria as a centre of Greek philosophy in later classical Antiquity. *Illinois Classical Studies* 18. 307-325.
- Busse, Adolf. 1887. *Porphirii Isagoge et in Aristotelis Categorias*. Berlin: Reimer. Commentarium (Commentaria in Aristotelem Graeca, IV-1).
- Busse, Adolf. 1895. *Ammonius in Categorias*. Berlin: Reimer. Commentarium (Commentaria in Aristotelem Graeca, IV-4).
- Busse, Adolf. 1898. *Ioannes Philoponus (olim Ammonius) in Categorias Commentarium*. Berlin: Reimer. Commentarium (Commentaria in Aristotelem Graeca, XIII-1).
- Busse, Adolf. 1900. *David [Elias] In Aristotelis Categorias Commentarium*. Berlin: Reimer. Commentarium (Commentaria in Aristotelem Graeca, XVIII-1).
- Busse, Adolf. 1902. *Olympiodori Prolegomena et in Categorias Commentarium*. Berlin: Reimer. Commentarium (Commentaria in Aristotelem Graeca, XII-1).
- Chase, Michael. 2003. *Simplicius On Aristotle. Categories 1-4. (Ancient Commentators on Aristotle)*, ed. by Richard Sorabji). London: Duckworth.

- Chriti, Maria. 2018. *Aristotle as a Name-giver: The Cognitive Aspect of his Newly Assigned Terms*. Center for Hellenic Studies: University of Harvard. <https://chs.harvard.edu/CHS/article/display/1304?menuId=181>.
- Chriti, Maria. 2019a. Aristotle's semiotic triangle in *On Interpretation* and the *Categories*. Language and thinking in the frame of logic. *Proceedings of the World Congress Aristotle 2400 years*, ed. by Demetra Sfendoni-Mentzou. Thessaloniki: AUTH-Interdisciplinary Centre for Aristotle Studies. 489-495.
- Chriti, Maria. 2019b. Neoplatonic word classes that designate Aristotle's *Categories*. *Blytyri* 8(1-2): 17-32.
- Chriti, Maria. 2021. Etymological proximities and onomastics: From Aristotle to Ammonius of Hermeias. *Greek Ancient and Medieval Etymology: Theory and Practice I*, ed. by Arnaud Zucker & Claire Lefevre. Berlin: DeGruyter. (Trends in Classics Suppl., 111). 37-51.
- Chriti, Maria. 2022. A new direction in Neoplatonic linguistics: Aristotle as an adherent of a 'specialist name-giver' by Ammonius of Hermeias. *Religions* 13(2) [*Conversion Debates in Hellenistic Philosophy and Early Christianity*, ed. by Eva Anagnostou, Georgios Steiris & Georgios Arabatzis]: <https://www.mdpi.com/2077-1444/13/2/172/htm>, accessed 03/14/2023.
- Ebbesen, Sten. 1981. *Commentators and Commentaries on Aristotle's Sophistici Elenchi*. Leiden: Brill. (Corpus Latinum Commentariorum in Aristotelem Graecorum 7, I-III).
- Ebbesen, Sten. 1990. Porphyry's legacy to logic: A reconstruction. *Aristotle Transformed*, ed. by Richard Sorabji. London: Duckworth. 141-171.
- Golitsis, Pantelis. 2008. *Les commentaires de Simplicius et de Jean Philopon à la Physique d'Aristote. Tradition et Innovation*. Berlin, NY: De Gruyter. (Commentaria in Aristotelem Graeca et Byzantina, 3).
- Hadot, Ilseraut. 1987. *Simplicius: savie, sonœuvre, sasurvie. Actesducolloqueinternational de Paris (28 Sept.-1 Oct. 1985)*. Berlin, NY: De Gruyter. Peripatoi, 15.
- Karamanolis, Giorgos. 2006. *Plato and Aristotle in Agreement? Platonists on Aristotle from Antiochus to Porphyry*. Oxford: Clarendon Press.
- Kotzia, Paraskevi. 1992. *Ο 'σκοπός' των 'Κατηγοριών' του Αριστοτέλη*. Thessaloniki: Aristotle University, Faculty of Philosophy.
- Lallot, Jean & Frédérique Idefonse, eds. 2002. *Catégories. Présentation, Traduction et Commentaires*. Paris: Éditions de Seuil.
- Praechter, Karl. 1922. Nikostratos der Platoniker. *Hermes* 57(4): 481-517.
- Saffrey, Henri-Dominique. 1989. Ammonios. *Dictionnaire des philosophes antiques*, dir par Richard Goulet, Vol. I, *Abam(m)on-Axiothea*. Paris: CNRS Éditions. 168-169.

- Sorabji, Richard. 1990. The Ancient commentators on Aristotle. *Aristotle Transformed*, ed. by Richard Sorabji. London: Duckworth. 1-30.
- Sorabji, Richard. 2005. *The Philosophy of the Commentators, 200-600 AD: A Sourcebook*. Ithaca, NY: Cornell University Press.
- Strange, Steven, K. 1992. *Porphyry: On Aristotle's Categories*, Ancient Commentators on Aristotle, ed. by Richard Sorabji. London, NY: Bloomsbury.
- Verrycken, Koenraad. 1990. The metaphysics of Ammonius son of Hermeias. *Aristotle Transformed*, ed. by Richard Sorabji. London: Duckworth. 199-232.
- Watts, Edward. 2004. Justinian, Malalas, and the end of Athenian philosophical teaching in AD 529. *Journal of Roman Studies* 94. 168-182.
- Westering, Leendert Gerrit. 1962. *Anonymous Prolegomena to Platonic Philosophy*. Amsterdam: North-Holland Publishing Company.
- Westering, Leendert Gerrit, Jean Trouillard & Alain Philippe Segonds. 2003 [1990]. *Prolégomènes à la philosophie de Platon*. Collections des Universités de France. Paris: Les Belles Lettres.

# Generative complexity and psycholinguistics: divorce American style<sup>1</sup>

**Jean-Michel Fortis**

Université Paris Cité and Université Sorbonne Nouvelle, CNRS, Laboratoire d'histoire des théories linguistiques, F-75013 Paris, France

---

## ABSTRACT

In the American scientific context of the 1950s, the confluence of information theory and behavioristic views seemed to hold the promise of a truly interdisciplinary psycholinguistics. However, the prospects opened up by this confluence were soon ruined by the advent of transformational grammar (TG). For reasons detailed in this paper, such was the attraction of TG that it became the nearly exclusive source of psychological hypotheses on linguistic processing. Correlating transformational complexity with measures of performance, such as response times, set the new methodological trend. Problems quickly crept in: psychologists could not solely rely on linguistic theory to account for their data and had to make room for heuristics and biases. Evidence for transformational theory was difficult to come by and TG was evolving at a pace psychologists had difficulty sustaining. On their part, linguists were reluctant to submit their constructs to psychological testing, so that the relationship of psychology to linguistics justifiably seemed to be one-sided. In all likelihood, countering the threat posed by psycholinguistics and defending the autonomy of linguistics underpinned Chomsky's affirmation that linguistic theory, with its own methods, lays a claim to psychological reality and does not need an auxiliary science.

---

1. My deepest thanks to John E. Joseph and Margaret Thomas for their comments, suggestions and corrections. It has been a privilege to benefit from the care of so diligent and expert reviewers.

Divorce was inevitable, but this short-lived episode of interdisciplinary research had significant consequences both in linguistics and in psychology. These consequences are dealt with in the last part of this paper.

**KEYWORDS**

history of linguistics, history of psychology, behaviorism, psycholinguistics, generative grammar, transformational grammar

**RÉSUMÉ**

Dans le contexte scientifique américain des années 1950, la confluence de la théorie de l'information et de conceptions behavioristes semblait promettre le développement d'une psycholinguistique véritablement interdisciplinaire. Toutefois, les perspectives ouvertes par cette confluence furent bientôt compromises par l'avènement de la grammaire transformationnelle (GT). Pour des raisons examinées dans cet article, l'attrait exercé par la GT fut tel qu'elle devint la source quasi exclusive d'hypothèses psychologiques sur le traitement du langage. La nouvelle méthodologie en vogue consista à corrélérer la complexité transformationnelle avec des mesures de la performance, comme par exemple les temps de réponse. Des problèmes surgirent rapidement : les psychologues ne pouvaient rendre compte de leurs données uniquement en se fondant sur la théorie linguistique et devaient prendre en considération des heuristiques ou des biais de traitement. Il se révélait difficile de fournir des preuves de la GT, laquelle évoluait à un rythme que les psychologues avaient du mal à suivre. De leur côté, les linguistes rechignaient à soumettre leurs constructions théoriques à l'épreuve de la psychologie de sorte que la relation entre les disciplines apparaissait déséquilibrée. Il est vraisemblable que la menace représentée par la psycholinguistique et le désir de préserver l'autonomie de la linguistique ont poussé Chomsky à soutenir que la linguistique peut prétendre établir la réalité psychologique de ses hypothèses par ses propres moyens et n'a pas besoin d'une science auxiliaire. Le divorce était inévitable, mais ce bref épisode de recherche interdisciplinaire ne fut pas sans conséquences, à la fois en linguistique et en psychologie. Ces conséquences sont l'objet de la dernière partie de cet article.

**MOTS-CLÉS**

histoire de la linguistique, histoire de la psychologie, behaviorisme, psycholinguistique, grammaire générative, grammaire transformationnelle

---

## 1. Introduction

It was not long after the term *psycholinguistics* came into wide use that the field experienced its first upheaval in the U.S. With the advent of transformational grammar (TG), says Reber (1987: 327), “the Chomskyan orientation captured the discipline to such an extent over the next two decades that the very name psycholinguistics became synonymous with the set of ideas that emanated on a regular basis from MIT”. In practice, this meant disrupting the “new interdisciplinary field” (Osgood & Sebeok 1965: xi), whose *modus videndi* rested on the association of behavioristic learning theories and information theory. This interdisciplinary framework may well have acted as a counter-model for generative grammar, especially when associated with an epistemology and a new psychology (cf. Fodor, Bever & Garrett 1974).

In this paper, we shall first examine the course of events leading to this upheaval in psycholinguistics and the appropriation of TG by psycholinguists. It will be seen that, once this appropriation was realized, a large part of psycholinguistic research in syntactic processing crucially involved measuring subjects’ performances against the amount of structural complexity posited by TG. The way this division of labor was conceived and the problems posed in particular by the evolution of TG will be examined.

Making psychological experiments the touchstone of linguistic theory was a potential threat to TG and to the autonomy of linguistics. Experimental testing, based in particular on transformational complexity, could disconfirm what was most distinctive about TG, i.e. transformations. How did linguists, and notably Chomsky, react to this threat and the encroachment of psychology? In tackling this question, we shall be drawn into a discussion of the complex of arguments marshalled by Chomsky to maintain the autonomy of linguistic theorizing. Or, to put it differently, it will be contended that some aspects of Chomsky’s metalinguistic reflection might be better understood as a self-protection device than as a coherent epistemology. Lastly, a few words will be said about the consequences of transformational psycholinguistics for the integration of processing

considerations into linguistic theorizing, and the shift of perspective this integration facilitated.

## 2. The historical context: the first happy marriage (or its prospect)

More than forty years separate Blumenthal's thin volume on the history of language psychology (Blumenthal 1980 [1970]) and the hefty compendium of Levelt (2013) on the same subject. Yet, if we believe Levelt, it would seem that the historical awareness of practitioners made little progress during these four decades. The notion that psycholinguistics emerged in the 1950 is still, says Levelt, "a widely shared opinion" (2013: 3), as it was when Blumenthal lamented the ignorance, on the part of his American colleagues, of an "illustrious earlier history of psycholinguistics" (1980 [1970]: viii).

Is this eclipsing due to a new experimental stance promoted from the 1950s on? This is quite doubtful: Osgood, who sanctioned the idea that the field was "new" (see the citation in our introduction) was well aware of the experimental past of behaviorist studies of language, including Soviet research (e.g. Osgood 1952). It is more likely that, for practitioners, the novelty lay in what they perceived as a genuine interdisciplinary collaboration between linguists and psychologists, a collaboration which was further legitimated by a new overarching technoscience, that is, information theory (IT).

At a time approximately coinciding with Osgood and Sebeok's *Psycholinguistics*, Hockett devoted a lengthy review to Shannon and Weaver's pioneering essay on information theory (Shannon & Weaver 1949). Hockett's text was actually more than a review: it was a discussion extrapolating the benefits linguistics could draw from looking at language in information-theoretic terms. Viewed from an engineering perspective, some specificities of natural languages could be brought out in a starker light, especially their massive redundancy on all levels and their corollary robustness. The identification of immediate constituents, a central issue after Bloomfield (cf. Wells 1947), could be reframed in terms of greater levels of indeterminacy ("entropy", in IT terms) at "cuts" between units, although

this criterion was quite probably not fool-proof, as Hockett himself remarked (1953: 87-8).

Rather surprisingly, Hockett makes no mention of Miller's *Language and Communication* (1963 [1951]), in which language design was also envisaged as an engineering question<sup>2</sup>. In particular, Miller gave quantitative estimations of the massive downsizing of theoretically possible discriminations in the acoustic channel, thus providing an assessment of the robustness of language which was much more detailed than the one given by Hockett. Miller also evoked transition probabilities between words as furnishing a realistic view of what speakers were doing when stringing forms together, and by so doing he showed how far he was from taking sufficient account of the hierarchical nature of linguistic structuring. By contrast, Hockett, as we just saw, was preoccupied with locating cuts in order to circumscribe units.

Hockett's tone in his review was also rather programmatic, while Miller's embrace of information theory and statistical reasoning evinced a good degree of confidence. For example, right after having mentioned his technique of probabilistic approximation to English strings, Miller stated that "the variety of sentence forms that a talker uses is not great, and probably the lengths of the patterns that are fitted together into these forms seldom exceed 10 words. The process of forging sentences is not inexplicable, and a clear formulation of what we need to know should lead to better observations and, eventually, better explanations of our verbal habits" (1963 [1951]: 192). What Chomsky, and Miller in his wake, would later declare as completely misguided was therefore deemed manageable at this point.

Whatever the discrepancies between Miller and Hockett, that a study as relevant as Miller's was not mentioned by Hockett may simply betray the fact that information theory percolated into psychology at a quite early date and before linguists took an interest in it. Miller had worked on signal intelligibility in a psycho-acoustic

---

2. This was not, however, the only perspective adopted by Miller. The book also contains chapters on personal styles, language acquisition, word-association protocols, the use of language in problem-solving and social patterns of communication. For a brief description of Miller's intellectual background, see Boden (2006: 286 ff.).

laboratory during the war, he was definitely attracted to mathematics and statistics, but he had strong reservations about the way some neo-behaviorists, like Hull and Spence, attempted to quantify behavioral variables (cf. Baars 1986: 303 ff. for a self-portrait of Miller). Obviously, this profile made him receptive to IT and ready to dissociate himself from neo-behaviorism, of which the Hullian brand was perceived as the most advanced until the early 1950s (Baars 1986: 60) and was still abundantly referred to into the 1960s (Leahey 1994: 207).

This bifurcation between linguistics and psychology in the reception of IT was not to last. The collective undertaking embodied in Osgood and Sebeok's *Psycholinguistics* could appear as foundational of a new interdisciplinarity, perhaps especially so in the eyes of Osgood, who bore the largest responsibility for the book's theoretical orientation (cf. Levelt 2013: 3-11 for a contextualization). We may observe, for instance, that the following introductory text to a chapter entitled "sequential psycholinguistics" is much in line with Osgood's later attempt to conciliate linguistic structure and probabilistic transitions between linguistic units (cf. Osgood 1963):

Study of the sequential or transitional structure of language behavior provides a meeting ground for linguists, information theorists, and learning theorists. The linguist, applying his own methods of analysis, discovers hierarchies of more and more inclusive units; the information theorist, usually starting with lower-level units such as letters or words, finds evidence for rather regular oscillations in transitional uncertainty in message sequences, the points of highest uncertainty often corresponding to unit boundaries as linguistically determined; and the learning theorist, working with notions like the habit-family hierarchy, finds it possible to make predictions about sequential psycholinguistic phenomena that can be tested with information theory techniques. (Osgood & Sebeok 1965: 93)<sup>3</sup>

The "meeting ground" referred to in this text rests on a consonance between the three disciplines in question. This kind of statement

---

3. The text is not signed, and it is unclear who penned the introduction to the chapter "sequential psycholinguistics", of which this text is excerpted. The rest of the chapter is due to several authors, including Floyd Lounsbury, a rather unexpected authorship since Lounsbury is better known for his involvement with structuralist ethnolinguistics.

(cited in Fodor, Bever & Garrett 1974: 53) would invite pro-generativist psychologists and linguists to lump together a bankrupt “taxonomic” linguistics with an outdated neo-behavioristic psychology (including a learning theory) and a lowbrow inductivism, using each to discredit the other. It is tempting to speculate that this antagonism helped them define a global endeavor whose ambition mirrors that of their opponents.

### 3. Situating Chomsky’s attack on behaviorism

Boden (2006: 285) rightly observes that “at base, Shannon’s theory—like behaviorism—was a beads-on-strings affair, for it concerned sequences of events conceptualized as Markov processes”, and this consonance implicitly exposed behaviorism to the attack launched by Chomsky (1957a) on associated explanatory initiatives, including probabilistic and Markovian accounts. However, the testimonies of Margaret Boden and George Mandler (Baars 1986: 263) agree on the fact that Chomsky’s critique of behaviorism really came to psychologists’ attention thanks to his arch-famous vitriolic review of Skinner’s *Verbal Behavior* (1959). It is not our purpose here to discuss the factors which led to the demise of behaviorism<sup>4</sup>. In this paper, our first concern lies in the role Chomsky and TG assumed in the evolution of psycholinguistics. However, the general context is certainly relevant for understanding the impact of Chomsky’s review, hence his status in the eyes of psychologists.

Testimonies (e.g. in Baars 1986) show that, in the 1950s, the Hullian behavioristic theory of learning was perceived as an empirical failure. Behaviorism at large was contested from several sides. In the United States, humanistic psychology was one of the paths taken by this critique of behaviorism’s dry operationism. Perhaps a sign of the theoretical “disorientation” of psychology diagnosed by

---

4. “Demise” may not be the best term. At any rate, it should not convey the idea that behaviorism has disappeared. Behaviorist protocols live on in cognitive psychology, which motivates Leahey’s assignment of methodological behaviorism and cognitive psychology to a super-category of “behavioralist” psychology (1994: 138-9). Behaviorist results (for example on the functioning of reinforcement) have been or are exploited in some areas of dynamic psychology, and today, in the algorithms of digital games and social networks.

Sigmund Koch as early as 1951 (Baars 1986: 67-68) was Skinner's perspective, which rejected speculation on inner variables and defended descriptivism, thus accomplishing a retreat to safe and arid empiricism. Other objections to behaviorism involving the role played by innate proclivities (or "instinct") and cerebral organization would soon surge, and some of these studies would later be marshalled by Chomsky in support of his nativism<sup>5</sup>.

Now, Skinner's fortes, experimental protocols and descriptions of stimulus-reinforcer-response contingencies, were missing from *Verbal Behavior*, although Chomsky's review made frequent mentions of animal conditioning and lever-pressing rats to ridicule the idea of making human speech a stimulus-dependent activity (for a discussion see Joseph 2002). To the eyes of a psychologist, the objection did not strike at the heart of Skinner's behaviorism, but could only concern a speculative book written in behaviorese, such as Weiss' rhapsody on behaviorist reductionism, which had so impressed Bloomfield and was non-experimental as well (Weiss 1929 [1925]). And in fact, Skinner's opus apparently fell on deaf ears: it seems that it got only one other review, globally negative, by a Soviet behaviorist in *Word* (Andresen 1991: 58). With hindsight, however, lumping together *Verbal Behavior* and animal learning theory appears to strike at the whole edifice. That is, Chomsky's critique targeted a book toward which most were indifferent, but its destructive power would soon be enhanced by the global obsolescence of behaviorism.

Chomsky's review left unscathed the synthesis between IT, learning theory and linguistics envisaged by Hockett and later promoted by Osgood and Sebeok's *Psycholinguistics*. Yet, this synthesis was linguistically better articulated than Skinner's account, and favorably considered by some linguists. However, he had already criticized this approach in his review of Hockett's *Manual of Phonology* (Chomsky 1957b), a text which may provide the missing link between *Syntactic Structures* and its attack on probabilistic processes, and the all-out

---

5. To the first category belongs the work of the Brelands on the "misbehavior of animals", that is, on responses that would be unexplainable if an organism were a tabula rasa and instinct not factored in (Breland & Breland 1961). To the second class belong studies of Hubel & Wiesel's famous studies (1963, 1965) on the innate organization of the visual cortex.

assault on the psychological branch of the “beads-on-strings” theory delivered in the review of Skinner’s *Verbal Behavior* (written shortly later, in 1958)<sup>6</sup>. It might be, as argued by Radick (2016), that the new mentalist, rationalist, nativist and anti-probabilist synthesis erected around TG took shape, in Chomsky’s case, in opposition to Hockett (or to the kind of unifying theory promoted by Osgood and Sebeok); to be sure it emerged progressively<sup>7</sup>. Whatever the case, the new synthesis invited psychologists to take a stance on linguistic theory and question their empiricism.

Finally, it should be noted that Chomsky’s opposition to behaviorism did not deter all behaviorists from seeking a compromise. Neobehaviorists like Osgood and Jenkins, who postulated internal processes mediating overt behavior, had already moved these mediational states from a peripheral level (as in Hull’s theory) to a more central, “cognitive” level. It would turn out that from their side, no strong objection to TG would be looming. In fact, both Jenkins and Osgood would show their readiness to endorse Chomsky’s criticism of Markov processes and to accept TG’s global framework, though criticizing linguists’ reluctance, including on the part of transformationalists, to delve into semantics (Jenkins & Palermo 1964; Osgood 1963). In spite of this good will, Chomsky seemed to ignore conciliatory solutions: they did not fit in with his definition of behaviorism (according to which this approach would not envisage systems “underlying behavior”, 1965: 193), and they were fundamentally “descriptivist”, in the sense of *Aspects*, i.e. based on the extraction of patterns and speech-habits (1965: 15). Indeed, in Palermo and Jenkins (1964), bona fide “empiricist” principles relying on the functional equivalence of stimuli and responses obtain general lexical classes and syntactic categories (such as SN and SV). This concern with a procedure used in constructing lexical and syntactic categories is in

---

6. A few years later, Fodor (1965) would mount an attack on neobehaviorist mediational accounts of meaning, specifically against the notion of mediating response, a central tenet of Osgood’s theory.

7. See Seuren (1998: 233-252), Koerner (2004: 40-41) and Radick (2016: 57), who points out that “there is not a whiff of antipathy to behaviorism” in the *Logical Structure of Linguistic Theory* nor in *Syntactic Structures*.

fact more in line with distributional accounts (e.g. Fries 1952) than with TG, in which it is swept under the rug.

#### 4. Miller and the introduction of TG into psychology

Miller was completely won over by the prospects opened up by the early Chomsky (cf. his testimony in Baars 1986: 207) and set out to acquaint psychologists with the theory, even if he was aware of the existence of alternative models which he described as equivalent (Miller 1962). Whatever the ultimate ground for Miller's choice of TG, his bet reflects TG's rise to prominence and victory over rival theories, for reasons well described in Nielsen (2010).

His introduction of TG into psychology, however, took an unexpected form, more comprehensible if we try to put ourselves in the shoes of a psychologist of this period. Firstly, TG was introduced in a collaborative book centering on goal-driven behavior (Miller, Galanter & Pribram 1960), not specifically on language<sup>8</sup>. This perspective reflected a long-standing concern of American psychologists working in one behavioristic perspective or the other: how can we account for apparently teleological behavior without resorting to mentalistic notions? (cf. e.g. Holt 1915; Tolman 1932). Teleological self-adjusting mechanisms (i.e. feedback mechanisms) had been on the market for some time, for example steam-engines regulators, but 'goal-directed' was associated with 'idea', or 'will' and 'will' with a ghostly faculty. When Bloomfield repudiated the mentalistic outlook, he scorned in particular its postulate of a "non-physical factor" called *will* (1933: 32)<sup>9</sup>. But in the 1950s, feedback mechanisms, like

---

8. On the background to this collaboration and the book that it spawned, see Boden (2006: 337). On the latter, a mischievous George Mandler commented that "Miller wrote it, Galanter takes credit in it, and Pribram believes it" (Boden, *ibid.*).

9. In *Language*, will and "mentalism" are opposed to "mechanism". A "mechanical" account makes no appeal to a "mind". Prima facie, mechanism fits in well with Bloomfield's conception of phonetic change, which he (like Neogrammarians) regards as driven by absolute regularities and independent of functional factors (for example, the fact for a phoneme, to bear a grammatical function). Yet, such regularities are not presented as "physical" (unlike in

anti-aircraft gun-control servomechanisms, were far more visible, and the reduction of purposive behavior to feedback mechanisms had already been proposed by Rosenblueth, Wiener, and Bigelow (1943).

Secondly, goal-directed behavior, envisaged as a feedback mechanism, could be connected with the concept of reflex and given a very general import through the idea, defended by Dewey (1896), that perception was fully integrated with sensori-motor coordination. Or to put it differently, the feedback loop involved in this integration could be envisaged as the basic behavioral unit (Miller, Galanter & Pribram 1960: 26-31). In its general pattern, this unit is very simple: the organism tests an input against some criteria, responds until there remains an incongruity between the result of the test and the criteria, and exits the loop when congruity obtains. The acronym *TOTE* (Test-Operate-Test-Exit) is the name chosen to designate this basic pattern (*ibid.*).

The behavior of an organism as simple as a tick cannot be modeled with a single TOTE unit. That is, TOTE units need to be sequentially and hierarchically arranged. Such an organization is called a *plan* and, most importantly, the call of subroutines by superior routines is assimilated to symbolic processes in a computation. On this account, the inner manipulation of symbols, including the high-level one traditionally designated as "inner speech", is the process whereby plans may be organized and reorganized. This is the important contribution that Miller *et al.* bring to the construction of cognitive psychology: several American psychologists had grappled with the inflexibility of stimulus-response chains, notably Tolman (1932), but symbolic computation based on TOTE units had not yet served to capture the fact

---

Neo-grammarians) but as a good heuristics (Bloomfield 1933: 357-358). In short, mechanism is not a plea for physicalist reductionism. In fact, behaviorism may appear as a scientifically acceptable alternative to such a reductionism, since, in the version of Weiss, it makes room for a "biosocial" level of analysis (in spite of Bloomfield's claim of a shared "physicalism" between Weiss and the Vienna School logical empiricists in Bloomfield 1936). The point is that Bloomfield's behaviorism and Chomsky's cognitivism may have something in common: finding a framework that does not entail an adoption of a lower-level reductionism (mechanism for Bloomfield, behaviorism for Chomsky) and safeguards the scientific credentials of linguistics.

that organisms may display flexibility, i.e. find a route alternative to the S-R chains they have been conditioned to.

The computer analogy, the existence of processing languages simulating hierarchical reasoning, and the distinction made between the abstract machine and its implementations entailed that complex hierarchical procedures could be posited without worrying too much about their physical realization. This legitimated the claim that TG deals with psychologically real processes, but need not concern itself with neurophysiology<sup>10</sup>. The coming of a new mentalism was heralded, and set in opposition to “Bloomfieldian” or “taxonomic” grammar (Katz 1964; Fodor, Bever & Garrett 1974). Yet, as we shall see, objections to taxonomic grammar on account of its association with an outdated (behaviorist) psychology will prove somewhat ironic from linguists who will end up insulating themselves from current psychology.

In the all-encompassing book of Miller *et al.*, and as befits a general psychological essay on behavior, language is introduced very late, after basic abilities and organic parameters (drives, needs, instincts, motor skills, memory) have been discussed. The confidence in IT expressed in Miller’s 1951 book has now evaporated and the rejection of the old synthesis is fully in line with Chomsky: Markov-type processes are rejected as unrealistic generators in view of the infinite linguistic capacity of human speakers. This capacity is distinguished from processing bottlenecks (such as memory limitations), following Yngve (and anticipating Chomsky; Miller, Galanter & Pribram 1960: 148). As for the reasons motivating the adoption of transformations, they conform to arguments presented in *Syntactic Structures*, in par-

---

10. “...it is clear that the linguist, though he claims that his theory describes a neurological mechanism, cannot immediately translate the theory into neurological terms, i.e. into talk about synapses, nerve fibers, and such. But – and this is the crucial point in showing that the mentalist is not a psychophysical dualist – this failure to have a ready neurological translation means only that he cannot yet specify what kind of physical realization of his theoretical description is inside the speaker’s head” (Katz 1964: 129). It is only required of a theory that it be “consistent” with psychology and neurophysiology (*ibid.*: 133). In this occurrence, Katz makes explicit reference to Putnam’s famous analogy between mental states and the logical states of a Turing machine (Putnam 1960), that is, to the view today identified as “(machine state) functionalism.”

ticular, regarding the economy of description<sup>11</sup>. But their appeal also lies in their being comparable to these symbolic processes which are hypothesized to underlie the reorganization of behavior and which Miller *et al.* associate with the rejection of S-R chains.

## 5. The appeal of a transformationalist psycholinguistics

An incentive for interpreting TG in a psychological way resided in Chomsky's referring to his theory as "mentalistic", i.e. "concerned with discovering a mental reality underlying actual behavior" (*Aspects*: 4). It should be noted too that Chomsky's proclaimed reliance on the native speaker's intuition could in itself be perceived as mentalistic, and as such, as a dubious introspective exercise<sup>12</sup>. This fact alone made him side with non-behaviorists. In addition, on Chomsky's part, there was apparently no reluctance to engage in an interdisciplinary enterprise: according to George Miller's testimony, Chomsky was far from averse to a psycholinguistic investigation of TG, at least initially (cf. Baars 1986: 243). In a chapter written in collaboration with Miller, the sketch of a processing model is drawn in broad strokes (Miller & Chomsky 1963). The model includes a "deep" processing stage whose output is the structural description of a sentence, complete with its transformational history. A major justification is that grammatical relations can only be recovered when this output is delivered (*ibid.*: 480).

On psychologists' side, generation and transformation were congenial to experimenters who thought in terms of dynamic processes, and had to measure response times (Greene 1972). Importantly, TG seemed to propose a division of labor between linguists and psychologists along the distinction between an account of "actual behav-

---

11. Miller *et al.* contend that greater economy is needed for learning a language in a "finite childhood" (Miller *et al.* 1960: 151). The non-commensurability between finite exposure and infinite creativity is the first feature to be integrated into Chomsky's ramified and complex argument in favor of nativism (cf. Thomas 2002 for a history).

12. For a telling rejoinder of Hill to Chomsky's use of introspective data, cf. Harris (1993: 54).

ior” and the “internalized grammar” which projects an infinite set of sentences (in the terms of Chomsky 1959), and, from 1963 on, between the elaboration of a grammar reflecting the “competence” of a speaker and a theory of “performance” (Léon & Riemer 2015: 131-7). On the other hand, it should be noted that Chomsky moved from a conception in which a grammar generates sentences and assigns degrees of grammaticalness to them, to a conception in which a grammar generates grammatical sentences only, and from there to a conception in which the set of grammatical sentences is obtained by factoring out all parameters related to performance (such as the limitation of memory span)<sup>13</sup>. Surely, psychologists (and, indeed, some linguists) may be inclined to resist the idea that memory span or various heuristics cannot enter into the very design of human language. We shall come back to this issue. And more to the present point, this increasingly abstract notion of grammaticalness may pull grammar out of the reach of psychologists, who would be entrusted with the task of studying its implementation in the mind/brain. The following passage suggests this interpretation:

When we speak of a grammar as generating a sentence with a certain structural description, we mean simply that the *grammar* assigns this structural description to the sentence. When we say that a sentence has a certain derivation with respect to a particular generative grammar, we say nothing about how the speaker or hearer might proceed, in some practical or efficient way, to construct such a derivation. These questions belong to the theory of language use – the *theory of performance*. (Chomsky 1965: 9, my italics)

To compound matters, this division of labor may not coincide fully with disciplinary lines. In the following pages, Chomsky proceeds to demonstrate how judgments of acceptability – for the average psychologist, a dubious introspective exercise – may serve to draw conclusions about the “perceptual device” used in parsing sentences (1965: 14-15)<sup>14</sup>. In this circumstance, the linguist is encroaching on

---

13. Chomsky used the expression “degrees of grammaticalness” before the distinction between acceptability and grammaticalness got established (cf. Léon & Riemer 2015).

14. A classification of sentences into “degrees of grammaticalness” was present as early as 1955 in *The Logical Structure of Linguistic Theory* (1975 [1955]), where

what would seem to be the psychologist's bailiwick, the theory of performance. All the same, this was also an encouragement to conduct psychological research in an interdisciplinary spirit but with psychology as *ancilla linguisticae*. As for the study of language use by itself, without the guidance of hypotheses on the underlying system, it was not deemed worthy of a "serious discipline" (1965: 4). In short, Chomsky made a distinction between a respectable theory of performance and the frivolous study of language use, the latter being probably a matter of determining why an utterance occurs in a given context (and encompassing, therefore, probabilistic accounts). The theory of performance can be identified with TG-inspired psycholinguistics and all conclusions derivable from judgments of acceptability.

## 6. Testing the psychological reality of transformations

Though TG-inspired psycholinguistics does not reduce to experiments specifically bearing on transformations, these experiments are crucial: transformations are a construct proper to TG (unlike constituent structure). Ultimately, as we shall see, results delivered by the relevant psychological variables (items recalled and response time) would not be found to be straightforwardly related to transformational complexity.

At first, however, there was hope that TG could provide a plausible blueprint for psychological processes. An early experiment by Miller & McKean (1964) seemed to show that response times were proportional to transformational complexity. Further, the results exhibited a neat additive pattern: a transformation was found to take an approximately equal time in different contexts, and performing two transformations yielded a response time close to the sum of the transformations when these were processed singly. Thus, transforming an affirmative active sentence into its passive counterpart, and an active negative one into its passive negative counterpart took approximately equal time. On the other hand, performing a negative

---

intuition is opposed to probability of occurrence, with Miller and Hockett in view (cf. Léon & Riemer 2015).

and a passive transformation on the same sentence yielded a time approximating the sum of the average response times obtained for each kind of transformation. For reasons we shall not go into here, other psychologists believed that transformational complexity would be better assessed by measuring the memory capacity taken up by their storage. This criterion was used by Savin & Perchonock (1965), who demonstrated that the more transformationally complex the sentence, the more it depleted the storage capacity of their subjects.

Several kinds of consequences arose when other parameters were found to interact with transformational complexity. First, some results could be interpreted as contradicting the additive hypothesis, or, alternatively, as entailing that subjects could bypass the transformational stage. For example, in an experiment conducted by Slobin (1966), subjects responded as quickly to passive sentences as to their active counterparts when the scene depicted in the passive sentence conformed to a familiar strong asymmetry between the participants (so called “non-reversible sentences”, e.g. *the flowers are being watered by the girl*, vs “reversible sentences”, e.g. *the boy was hit by the girl*). More striking even was Fodor & Garrett’s observation (1967) that transformationally more complex sentences could be decoded more easily than simpler ones. Thus, *The first shot the tired soldier the mosquito bit fired missed* was actually easier to understand than the same sentence without the adjectives (i.e. a less complex sentence), presumably due to the presence of semantic cues.

The central role of processing complexity casts doubt on the psychological relevance of one of the original motivations of transformation, namely that they provide an economical way of generating related sentences. This can be exemplified by Dryer’s typological study on the position of sentential NPs within their matrix clause (1980), in which Dryer takes stock of previous studies published in the 1970s and submits his own account in terms of positional tendencies. If, for reasons we shall not go into here, *That Fred loves Mary is obvious* is more difficult than its variant with an extraposed sentential NP (*It is obvious that Fred loves Mary*), why would subjects bother to generate the latter utterance from a more difficult structure corresponding to the former sentence? If economy of linguistic description *must* have psychological relevance, as argued by Chomsky, it

cannot but conflict with ease of processing in cases like these. Either economy of description supersedes considerations of processing costs, and all psychological and typological results are of secondary importance, or economy of description simply is a meta-theoretical criterion which has nothing to do with “psychological reality”. But in the eyes of psychological realists like Chomsky, the latter conclusion is ruled out.

As we have seen, data were often too messy to be accommodated by a simple mapping from transformational complexity to processing time. In some cases, this gave rise to a second kind of consequences: psychologists felt compelled to supplement transformations with the decision procedures needed to account for their results. Negative sentences proved to be a very delicate matter: interpreting the pattern of results entailed that the truth-value of a negative sentence had to be factored in<sup>15</sup>. Typically, response times to true negative sentences were longer than to false negative ones (for a review, see Greene 1972). The decision procedures designed to capture the patterns of results took the form of flow charts in which what remained of TG was the assumption that kernel sentences furnished the canonical representation of sentences as bearers of truth-value, with deviations requiring additional processing steps (Trabasso *et al.* 1971; Chase & Clark 1972). In addition, a natural motivation for the advantage of false negatives over true negatives presented itself: on the assumption that negations are most naturally used for denying a state of affairs held as true, false negatives are closer to “naturalistic” contexts (cf. Greene 1972; Wason 1972). The functional motivation invoked in this occurrence by psychologists like Wason was quite extraneous to TG’s concerns. Further, a connection was thereby established between linguistic comprehension and “natural”, not necessarily normatively logical, reasoning. Deductive biases in natural reasoning were not a new concern in themselves, since philosophy was born partly out of a reaction to sophistry. The novelty lay

---

15. Negative sentences are transformationally more complex than kernel sentences in the framework of *Syntactic Structures*. This is not so in *Aspects*. The problem remains, however: for psychologists, the interaction between truth-value and negation still has to be accounted for, peripheral as it might be for linguists.

in the systematic experimental investigation of hypothesis formulation and logical deduction (one of Wason's pet subjects). This issue, then, takes us to a different continent. Within generative grammar, "natural" reasoning would be broached by supporters of generative semantics only (cf. Lakoff 1970).

The varying status of negative sentences from *Syntactic Structures* to *Aspects* illustrates yet another kind of consequence brought about by the wedding of psycholinguistics to TG: the interpretation of performance data hinged on the current state of the theory. Thus, it was countered that in Miller & McKean's experiment cited above, transformations had not been counted properly: negation also implied the *do*-support transformation for non-auxiliary verbs<sup>16</sup>. The results, therefore, were not valid. But in the model of *Aspects*, Miller and McKean's results were defensible again: the *do*-support transformation was post-deep structure, and Miller & McKean's results could be taken to measure the relative complexity of deep structures (DS), not of transformations (cf. Fodor, Bever & Garrett 1974: 235-241). Likewise, experiments on the role of deep-structure frequency in recall depended, for their interpretation, on the hypothesis that pronouns are inserted post-DS, that is, that they replace full nominal phrases present at DS, but this was controversial (Fodor, Bever & Garrett, *ibid.*: 259-263). Not only did psychologists have to look to linguistics for their hypotheses, but the interpretation of their results depended on the current state of generative theory, and this was rapidly changing.

To sum up, elaborating a psychologically plausible theory of performance required making room for semantic factors, decision procedures and functional considerations. Further, it was not always clear how to reconcile psychological processing factors with an a priori criterion such as economy of description. More circumstantially, for psychologists to keep up with the pace of theoretical change in linguistics, experiments had to be framed in the terms of the latest version of TG, and results interpreted or reinterpreted in the light of this version. This put a double strain on psychologists: that of separating the grammatical core from factors linguistic theory regarded

---

16. This issue and others related to the proper way of counting transformations are raised by Fodor, Bever & Garrett (1974: 231-234).

as extraneous, and of updating their hypotheses in conformity with theoretical changes made by linguists.

## 7. Some pills psychologists found hard to swallow

Ideally, the division of labor between linguistics and psychology should yield results that are complementary and cross-fertilize both disciplines. However, this division also entails a methodological parting of the ways: linguists who make a distinction between 'grammatical' and 'acceptable' filter out some acceptable sentences, and accept as grammatical unacceptable ones, on the ground of their conformity to generalizations they believe to have gained. For psychologists, actual data (abstracting away from statistical corrections, aberrant responses etc.) are all *bona fide*, and the identification of performance factors is a complex empirical matter. On the other hand, unattested or very unlikely sentences which linguists regard as grammatical can hardly figure in psychologists' empirical evidence.

Chomsky would come to depict the generativist attitude as the sound scientific practice of a "Galilean" science, a quest for principles idealizing away from theoretically irrelevant aspects of natural phenomena (for a critique, see Riemer 2009). Further, the formulation of principles being the linguist's bailiwick, nothing warrants that the data and processing hypotheses submitted by experimenters are relevant to linguistic theorizing. This situation was apt to make linguistics impervious to psycholinguistics' meddling in its affairs, and could be perceived as leading to a form of linguistic isolationism. The risk was clearly articulated by Greene: "the danger is that any data fitting in with a linguist's preconceptions about specifically linguistic universals will be counted as evidence for the basic structure of language, while any awkward data can be put down to the influence of functional 'performance' factors" (Greene 1972: 104)<sup>17</sup>.

---

17. A suspicion that is not unwarranted when we read: "A grammar is simply an axiomatic representation of an infinite set of structural descriptions, and the internal evidence in favor of the structural descriptions modern gram-

The term “preconceptions” seems to reflect the perspective of an agnostic observer, in a time when linguistic theories had not only failed to reach any consensus (and have not to this day) but, in the generative camp, were involved in a bitter strife pitting the syntax-centred interpretive stream against the generative semantics movement (Newmeyer 1986; Harris 1993; Huck & Goldsmith 1995; Fortis 2015). As a matter of fact, the peak of “generative” psycholinguistics corresponds approximately to the period during which generative semantics is in full swing, roughly between 1965 and 1975<sup>18</sup>. The discord could not but shake psychologists’ confidence in the bet they had placed on TG. Fillenbaum (1971) gives voice to this disorientation: even leaving aside Yngve’s phrase structure model and Lamb’s stratificational theory, he asks, should psycholinguistics remain faithful to standard deep structures, or should it espouse generative semantics? And what should psychologists make of the new relevance attributed to surface structures for semantic interpretation, and wielded by Jackendoff (1969) and Chomsky (1972a) against generative semantics? Whereas the most convincing experimental results had been secured for (deep) structures, not for processes (i.e. transformations, cf. Fodor, Bever & Garrett 1974), controversy over the very nature of deep structures was now rife and the alternative views put forth by generative semantics were not without appeal. Albeit generative semantics largely handled semantics with syntactic tools, their theory was no longer syntax-centred but made proposals on the ways semantic representations, or ‘thoughts’ (the term of Lakoff 1976 [1963]), could be mapped to forms, and thus promised a more integrated blueprint of the speaker’s articulation of content. Chomsky’s rebuttal (1972a: 69 ff) that the Standard Theory is not only interpretive but, taken in the opposite direction, is also a the-

---

mars generate is so strong, that it is difficult to imagine their succumbing to any purely experimental disconfirmation” (Fodor & Garrett 1966: 152).

18. Reber (1987: 328) provides the *Science Citation Index* for “Chomsky” from 1964 through 1983 (I assume the discipline concerned is psychology but Reber is unclear on that point). A sharp decline is observable after 1976. This decline occurs in the context of a burgeoning of publications in AI, cognitive psychology, neurosciences etc. In this context, and apart from the demise of generative psycholinguistics, we may surmise that a number of psycholinguists have turned to other frameworks (see below, section 10).

ory of production, or that Chafe's and Fillmore's semantics-centred accounts are notational variants of his own, were hardly more than a sleight of hand and could scarcely mask the lack of consensus which prevailed in linguistics. As appealing as their views might have been, however, generative semanticists were not offering a unified theory, but rather scattered descriptions tackling all kinds of fields newly absorbed into linguistics: quantification, focus, belief-contexts, decompositional semantics and adverbial scope, possible worlds and presuppositions... The over-ambitious scope of their inquiries and a hectic pace of theoretical evolution obviously did not facilitate the elaboration of a unified account and, although a textbook was announced, it failed to materialize (Harris 1993: 219). On the other hand, textbooks of standard TG were being churned out (Bach 1964; Roberts 1964; Thomas 1965; Jacobs & Rosenbaum 1968; the list is not exhaustive), and the institutional clout of interpretive generativists was stronger (Harris 1993).

Eventually, the domains covered by generative semantics were dispersed among specialized disciplines, sometimes incorporated into mainstream generative grammar, and partly reelaborated into strands of cognitive linguistics. For example, problems related to quantification and which had proved to be thorny for Standard Theory were reformulated in formal brands of semantics. From her own testimony, we learn that the problems attending quantifiers were the main reason why Barbara Partee turned away from TG and to Montague's theory (Partee 2011). The use of variables and formulas akin to predicate calculus introduced by generative semantics was imported into generative grammar (in the module of "Logical Form"). Questions involving belief-contexts, possible worlds and presuppositions, then much debated in philosophy and linguistics, would be appropriated in particular by a theory affiliated to cognitive linguistics, the mental spaces approach proposed by Fauconnier, starting in 1978 (see Fortis & Col 2018). More globally, generative semanticists mutated into cognitive linguists (approximately after 1975) and redirected their semantic inquiries within this new theoretical network (Fortis 2015).

The state of dissent among linguists was thus accompanied by a fragmentation of semantics into distinct approaches. For some psy-

chologists and linguists, this situation was compounded by doubts about the possibility of identifying psychologically real semantic representations with componential structures assumed to underlie generation. For example, Fodor et al. (1975), partly from experimental evidence, argued against the psychological reality of definitions assumed by generative semanticists to underlie some lexical items, a classic example of which was the analysis of *kill* as derived from 'cause to become not alive' via a lexical transformation (Predicate Raising, cf. McCawley 1968, 1971). This, again, militated against correlating complexity of linguistic constructs with complexity of processing. Meanwhile, psychologists were drifting to non-behaviorist, constructivist approaches which emphasized that perception, memory, categorization and semantic processing involve recoding, reanalyzing and transforming the sensory input subjects are exposed to (Neisser 1967). Bransford and his associates insisted that the comprehension of sentences could not be based solely on a linguistic analysis treating them as self-contained objects; that is, semantic representations rest on knowledge of the world and the abstraction of "wholistic ideas" summing up various formulations of the same situation (Bransford & Franks 1971; Bransford & Johnson 1973). The constructivist perspective thus tended to divorce semantic representations from any kind of deep structure specifically directed at generating surface sentences.

To sum up this discussion, in addition to the experimental difficulties outlined above, a number of circumstances negatively affected the collaboration of psycholinguistics with generative grammar: the one-sidedness of psychology's relationship with linguistics, the inner strife shaking TG and the ensuing theoretical changes, the fragmentation of semantics into sub-fields taken over by various disciplines, and the new constructivist perspective on semantic representations in psychology.

## 8. "Psychological reality" and the autonomy of linguistics

Given circumstances adverse to a balanced collaboration between psychology and linguistics, how should we interpret the statement

made by Chomsky in *Language and Mind* (1972b: 1), and repeated afterwards, that linguistics “is a branch of cognitive psychology”? In *Rules and Representations* (1980), written after the divorce between transformational psycholinguistics and GG, the answer is quite plain: linguistics, by its own methods and criteria of validation, is a psychological discipline, more precisely a branch of theoretical psychology “primarily concerned with the genetically determined program that specifies the range of possible grammars for human languages and the particular realizations of this schematism that arise under given conditions” (*ibid.*: 202).

What Chomsky rejects is the view that “certain types of evidence are held to relate to psychological reality, specifically, evidence deriving from studies of reaction time, recognition, recall etc.,” while theoretical explanations advanced to account for evidence based on informant judgments would have no claim to “psychological reality” (1980: 192). There is no such epistemological divide, he says:

...any theory of language, grammar, or whatever carries a truth claim if it is serious, though the supporting argument is, and must be, inconclusive. We will always search for more evidence and for deeper understanding of given evidence which also may lead to change of theory. What the best evidence is depends on the state of the field. The best evidence may be provided by as yet unexplained facts drawn from the language being studied, or from similar facts about other languages, or from psycholinguistic experiments, or from clinical studies of language disability, or from neurology, or from innumerable other sources. [...] But there is no distinction of epistemological category. In each case, we have evidence – good or bad, convincing or not – as to the truth of the theories we are constructing, or if one prefers, as to their “psychological reality”, though this term is best abandoned, as seriously misleading. (Chomsky 1980: 109)

This is a strange argument. What counts as legitimate “evidence” each time hinges on the criteria and on the methods which prevail within a given discipline. Even if we assume that the evidential base of GG is completely sound (but see Riemer 2009), a well-supported generative description simply does not comply with the requirements of experimental psychology. In the eyes of a psychologist, such description must await confirmation and cannot be regarded as valid simply because it is impeccably argued for from a linguistic point

of view. By the same token, and to reiterate a point already made above, “evidence” is shaped by the methods employed to obtain it: reaction times in a psychological experiment are part of what counts as evidence, hence it is requested from the theory that it account for the pattern of reaction times. In practice, this means that parameters extraneous to linguistic forms and their well-formedness must be factored in, for example the truth-value of negative transforms (see above, section 6). This may cause (and has caused) psychological models to diverge from linguistic ones, since psychological models must be specifically designed to fit behavioral data.

A possible way to understand Chomsky’s argument, I suggest, is to consider it in its historical context, as a strategy put in place after the demise of TG and elaborated with the goal of protecting the autonomy of linguistics. In effect, the argument keeps psycholinguists from meddling in the linguist’s affairs, but does not in the least relinquish linguistics’ claim to psychological reality.

What we interpret here as a disguised plea for autonomy is well illustrated in the text in which it is put forth. In the chapter of *Rules and Representations* (1980) entitled “On the biological basis of language capacities”, Chomsky pays lip service to the sum of neurophysiological knowledge in the book which has inspired his own chapter title, Lenneberg’s *Biological Foundations of Language* (1967). Besides a nod of approval to Lenneberg’s argument for innate mechanisms or predispositions, no attempt is made at taking stock of the “evidence” in neurophysiology, aphasiology, language acquisition, etc. Instead, Chomsky devotes a few pages to demonstrating the generalization of the *wh*-island constraint to sentences of the type *I found a violin to play sonatas on*, a generalization he sees as a ground for asserting the soundness of the *wh*-island constraint, and, as a consequence, its “psychological reality” (1980: 193-197)<sup>19</sup>.

While the point made by the generalization of the *wh*-constraint concerns the *description* of linguistic structures, the autonomy of linguistic argumentation seems more difficult to maintain when we are

---

19. The *wh*-island constraint, originally put forth by Ross, prohibits *wh*-movement from a *wh*-clause. For example, \*Which sonatas did you find a violin to play on is ruled out by assuming an underlying *wh*-clause of the form *I found* [<sub>NP</sub> a violin [<sub>S</sub> which for PRO to play sonatas on *t*]].

dealing with language *acquisition*. Yet, the arguments marshalled by Chomsky in favor of his nativist stance essentially rest on invented examples, not, for instance, on a corpus of children's utterances. His objection to analogy as a source of linguistic creativity, for instance, merely rests on the possibility of false analogies between a couple of sentences created for the purpose of giving an illustration of a misleading formal parallelism<sup>20</sup>. The objection, however, is weakened by real-life observations, which tend to show that children learn constructions in a piecemeal fashion, and extract patterns very progressively, as has been shown by Tomasello (2003)<sup>21</sup>. False analogies of the kind mentioned by Chomsky involve a generalization over constructions of quite different types, a step that is in need of evidential support and would be *prima facie* unlikely if we extrapolate from Tomasello's observations. It is beyond the scope of this paper to examine Chomsky's arguments for his brand of nativism. Suffice it to note here that they are by and large independent from acquisition studies, hence fall in line with the autonomy proclaimed in *Rules and Representations*, and that they form a mixed bag of a priori reasoning (e.g. on constraints on possible hypotheses made by children) and statements in need of empirical support (as in the just mentioned example). In other words, the autonomy of linguistics may verge on speculative thinking in matters regarded elsewhere (in the psychology of acquisition) as empirical.

---

20. Cf. *Knowledge of Language* (Chomsky 1986: 8, 105). Suppose analogy were the parsing procedure used in analyzing sentences. Then (a') and (b'), which stand in analogous relation to (a) and (b), would be parsed identically:

(a) John ate an apple.

(a') John ate.

(b) John is too stubborn to talk to Bill.

(b') John is too stubborn to talk to.

But they are not parsed identically. *Ergo*, analogy is not the procedure used.

21. An objection already raised by Prideaux (1975) when discussing the transformational account of children's acquisition of interrogative sentences put forward by Klima & Bellugi (1966).

## 9. The linguistic aftermath

The brief marriage of transformational psycholinguistics and generative grammar brought into relief processing factors involved in the parsing of structures. This offered a favorable context for an inversion of perspective, the idea that processing demands and heuristics play a role in shaping linguistic structures themselves. In this new perspective, complexity of processing is no longer a property of the mechanism operating on independently motivated linguistic structures; rather, linguistic structures partly reflect strategies aimed at reducing processing complexity. This inversion seems to have occurred in several steps, and through different ways. The studies mentioned below are meant to illustrate this evolution, but others might have been cited.

An early step in this change of perspective is the study of Fodor & Garrett (1967). Their reasoning is worth retracing. We have reasons to believe, they point out, that understanding a sentence is not straightforwardly related to transformational complexity (see section 6). Processing rests on cues of various kinds which together help subjects home in on the right analysis. An example of a cue which helps reduce parsing uncertainty is furnished in the initial string of (1). Uncertainty is relatively high in (1), since it could be followed by a conjunct or the predicate of a relative clause or a relative clause, etc. The addition of *whom* in (2) rules out one possibility, that is, the conjunct, thus reducing processing complexity.

(1) *The boy the girl...* [*The boy the girl...and the man left...*, *The boy the girl...met.....*]

(2) *The boy whom the girl...*

Note that *wanted* in (3) would augment uncertainty with respect to *met*:

(3) *The boy whom the girl wanted...* [or wanted to meet..., or wanted the man to meet...] vs *The boy whom the girl met*.

Uncertainty rises with *wanted*, because of its having alternate constructional possibilities. In this framework, then, subjects do not undo transformations. They use whatever cues are available along the

incoming string that is being processed, be they semantic, syntactic, lexical (e.g. the subcategorization frame of a verb).

A further step is taken by Bever in his paper on “the cognitive basis of linguistic structures” (1970), in which it is contended that processing heuristics (in his terms, “perceptual strategies”) reflecting cognitive biases (“perceptual constraints”) may *determine* the grammaticality of sentences. In his own words, Bever suggests “that certain aspects of sentence structure reflect the perceptual constraints placed on it by the child as he learns the structure and by the adult as he uses the structure” (1970: 313). One of the cognitive biases identified by Bever would be the default segmentation of strings into ‘actor action object...modifier’ roles, used by children acquiring a first language. This early bias would underlie a default perceptual strategy favoring in adults as well the analysis of an initial string (in English, a NP VP string, for example) as a main clause. A consequence of this perceptual strategy would be the introduction of a marker of subordination to prevent an initial NP VP structure from being interpreted as a main clause<sup>22</sup>:

(4) *The fact that / That / The fact the door was discovered to be unlocked amazed the tenants.*

But not

(5) *\*The door was discovered to be unlocked amazed the tenants.*

In Bever’s paper, other “aspects of sentence structure” for which “perceptual strategies” determine grammaticality involve pronominal coreference and constraints on the deletion of relative pronouns,

---

22. This strategy would be strong for English, but presumably weaker or non-existent in languages (like Japanese) which allow initial subordinate clauses with no initial marker of subordination. Note that the Main Clause strategy, if extended to all languages, would make processing left-branching structures (as, again, in Japanese) inherently more complex: whether a clause is a matrix clause or a subordinate one is undetermined until a complementizer or the clause second boundary is received (see the discussion in Dryer 1980). The “actor action object” bias in the linguistic construal of states of affairs and events has a long history (see e.g. Knobloch 1988: 357). The relative novelty of Bever’s account lies in the ontogenetic (rather than phylogenetic) role of this heuristic, and its operating as a default parsing strategy.

that is, issues that generative grammar handled in formal, configurational terms.

Among the attempts to derive structural facts from processing considerations, one particularly stands out for its typological scope and its ambition to explain observational generalizations, and that is the line of research concerned with word order universals. Perhaps the most remarkable efforts in this direction are due to Hawkins, who himself based much of his early theory (in Hawkins 1983) on Greenberg's essentially inductive universals (1966), and on Vennemann's (notably 1974a, 1974b) and Keenan's (e.g. 1973, 1978, 1979) tentative accounts to formulate very general principles constraining the alignment of syntactic order, for example head-modifier vs modifier-head, across different syntactic categories. In Hawkins' own account, one of the most important explanatory principles, the "Heaviness Hierarchy", ranks modifiers in their potential to move to the right of the modified element, eventually in violation of the dominant order of a given language. On Hawkins' view, "heaviness" acts as a psychological parameter: the term sums up a number of processing constraints which directly condition the syntactic order of constituents. With a typological scope, hence a cross-linguistic generality which far exceeds Bever (1970), processing costs are no longer seen as a filter on structures compatible with competence, but as "one of the causal factors explaining why languages employ the rules and structures that they do" (Hawkins 1983: 104).

That processing factors may account for facts of linguistic structure is also a point made repeatedly by Givón in his first book, *On Understanding Grammar* (1979). Givón's undertaking, however, goes beyond this. It illustrates the broader view that linguistic structure is functionally motivated, the reduction of processing costs being but one of the functions performed by grammar. A clear connection is thereby established between "psycholinguistic" considerations, and a broader functional perspective for which processing factors are one of the motivations accounting for grammatical structure, understood as a "solidification" of discourse structure, world categorization, cognitive and ontogenetic constraints, peppered with vestiges of diachronic evolution (1979: 3-4). Givón's attitude evinces a frustration we find expressed in the work of other scholars: generative grammar

merely stipulates rules which are descriptive of grammatical facts but which are not explanatory. The same quest for explanations external to formal generation characterizes one of Langacker's first "pre-cognitive" attempts. In his 1974 paper on raising rules, their very existence is justified by the higher "degree of prominence" which they assign to the raised constituent. Here, "prominence" appears to be a precursor of Langacker's future cognitive and monological conception of Gestalt-like saliency; it is not intended as a pragmatic notion. On the other hand, pragmatic considerations are at the forefront in Hooper [= Joan Bybee] & Thompson's study on "root transformations" (1973)<sup>23</sup>. Just like for Bolinger (1977) and others, for Hooper & Thompson the actual non-synonymy of transforms points to the semantic laxity of transformationalists. It is also an invitation to seek non-syntactic parameters explaining limitations on the productivity of transformations. In the case at hand, they hypothesize the motivation of root transformations resides in the speaker's intention to "emphasize some particular element", with the corollary that this element should not be presupposed in the discourse situation. Taking into account pragmatic intention introduces another layer of complexity to constraints on transformations, potentially interacting with processing factors. For example, the difference between (5) and (6) could be claimed to rest on the fact that the transformation Comp Preposing is ruled out when the main clause is asserted (as in example 5); in (6), *he said* is not the focus of the assertion but is claimed to be merely parenthetical.

(5) \**It's just started to rain, he didn't say.* [Comp Preposing]

(6) *It's just started to rain, he said* [parenthetical].<sup>24</sup>

---

23. Roughly, transformations in which the moved constituent is immediately dominated by S. A simple example is VP-preposing, illustrated for instance by *Mary plans for John to marry her, and marry her he will.*

24. In the first version of this paper, the examples taken from Hooper & Thompson concerned Negative Constituent Preposing. They were:

(5') \**That never in his life has he had to borrow money is true.*

(6') *It's true* [= parenthetical] *that never in his life has he had to borrow money.*

Both reviewers found (5') perfectly acceptable. John Joseph (p. c.) remarks that a number of sentences marked as ungrammatical by Hooper and Thompson are quite acceptable to him. This raises again the issue of the soundness of data

In these instances, issues mapped out by generativists and first tackled as a matter of formal grammar (in linguistics) and various heuristics (in psycholinguistics) usher into broader functionalist views. On the side of early functionalists and cognitive linguists, normal description is no longer deemed as explanatory and it is seriously doubted whether sentences related by transformations are cognitively, semantically or pragmatically equivalent.

## 10. By way of conclusion: some notes on the subsequent evolution of psycholinguistics

As we have just seen, the failed partnership of TG and psycholinguistics proved indirectly fertile by directing the attention to parameters involving functional motivation, cognitive, semantic and pragmatic constraints. The divorce, therefore, should not be evaluated too negatively. In the short-term, however, psycholinguists were left without a general blueprint for the syntactic processing of surface structures. As Frazier notes (1988: 15): “Considerable psychological evidence showed that the structures the grammar attributes to sentences are psychologically real, even if the transformational rules of the 1960-style grammar were not. Since the only known characterization of surface structures was the one provided by transformational grammar, this state of affairs presented an apparent contradiction. In response to this situation, several models of comprehension were developed in which grammatical information was either not used, used haphazardly, along with sundry heuristics and heavy reliance on lexical and world knowledge...” Clark & Clark’s psycholinguistics textbook (1977) is a good illustration: constituency (“the structures the grammar attributes to sentences”, in Frazier’s words) remains the firmest gain of experimental research and parsing has dissolved into a set of “strategies”, often amounting to paraphrasing grammatical and pragmatic facts as a set of instructions (“when you have a determiner, open an NP and look for a closing N”, “look for given information to precede new information”). This bottom-up, category-based tendency also manifests itself in their nod

---

based on the linguist’s intuition, a problem that seems to be particularly acute in generative grammar (Riemsdijk 2009).

of approval to lexically- and valency-based accounts (invoking for example Schank 1972). In a framework such as this one, complexity of processing can be calculated by the degree of deviation from the set of heuristics brought to bear on an utterance. However, since strategies encapsulate all sorts of structural facts, as well as semantic and pragmatic biases, we are no longer on the safe ground of a self-contained syntax, and the possibility of giving an operational definition of complexity, one amenable to experimental testing, appears as an unrealistic hope.

The divorce that took place in the 1970s is associated with the redirection of psycholinguistics toward low-level processes (e.g. word recognition, especially after Morton 1969), and toward the study of meaning (Tanenhaus 1988). While semantics was the subject of a growing interest on the part of American linguists, phonological and morpho-syntactic matters remained the center of gravity of linguistic theorizing. As far as linguistic semantics was concerned, and although the psychological reality of semantic constructs seemed to be the concern of every linguist, including supporters of formal models (e.g. Partee 1979), this preoccupation found very little echo in psychology. If we except a few scattered attempts (e.g. Fodor *et al.* 1975 on the psychological reality of definitions), the psycholinguistics of meaning and modelling systems with a psychological ambition do not serve as an ancilla linguisticae, as is attested, for example, by the flourishing of semantic descriptions of lexical items, texts and real-world knowledge in the form of predicate calculus and propositional systems (e.g. Kintsch 1974, Miller & Johnson-Laird 1976), and semantic networks (especially after Quillian 1966; for a critical overview, see Rastier 2010 [1991]). To a certain extent, the tables are turned: the redirection of psycholinguistics toward semantic questions opens up new niches of investigation for linguists. As argued by Honeck (1980), for example, Lakoff and Johnson's conceptual metaphor theory (Lakoff & Johnson 1980) should be situated in the context of an interest in metaphor on the part of psychologists during the 1970s, an interest which Honeck regards as a side-effect of the demise of transformational psycholinguistics. Likewise, prototype theory, predominantly rooted in psychology, would be instrumental in reviving lexical semantics in the U.S. (Fortis 2018). But note that in both fields,

the ties of linguistics with psychological and neuroscientific research, initially at least, were thin. Thus, after a simplified version of prototype theory had been imported into linguistics, linguists lost interest in the evolution of psychological theories of categorization, and very little experimental research was done. More significant interdisciplinary connections, however, would be progressively established in the field of conceptual metaphor, and increasingly so thanks to an intersection between the empiricist assumptions of this theory and pleas being made for a reevaluation of the role of the body in cognition from the 1990s on<sup>25</sup>. Today, this intersection is quite manifest in research conducted in the rather vast domain that goes by the name of “grounded cognition” (e.g. Barsalou 2008).

On the syntactic front and mostly after the 1970s, we can discern a turn to a conception of processing in which syntax forms a self-contained system again. This turn seems to have two main motivations. First, experiments showed that syntactic parsing strategies were applied in an “automatic” way, in particular immune from semantic biases. Early studies in this direction were able to cast doubt on what seemed to be a particularly robust manifestation of such a semantic bias, namely the role of argument reversibility in passives (see *supra*, section 6). By teasing out the effect of reversibility and the inherent plausibility of situations depicted in carefully controlled sentences, Forster & Olbrei (1973) showed that processing time was a function of syntactic complexity, whether the sentences were reversible or not. In addition, a whole literature on “garden-path sentences” and their online treatment developed out of an interest in the default principles used in building up syntactic structure (for an overview, see Frazier & Clifton 1996, introduction). Through this avenue of research, online parsing strategies exploiting purely syntactic principles offered the dynamic view of constituent structure that psycholinguists had been so glad to seize upon in the time of TG. The second motivation for regarding the syntactic processor as a self-contained system lay in the evolution of linguistic theory itself. This aspect is clearly articulated in Frazier’s defense of her Garden Path Theory of syntactic processing: her conception of grammar is modular in the

---

25. E.g. Varela *et al.* 1991; Lakoff & Johnson 1999; Joseph 2018 for a long-term historical account.

sense in which Government and Binding Theory is, assuming for example a separate processor for thematic relations. The irrepressible nature of the processing of linguistic form was also adduced by Fodor in support of his modular conception of mental architecture, in his *Modularity of Mind* (1983). Against any view arguing for an influence of contextual content on parsing, he declared rather tersely that “linguistic form recognition can’t be context-driven because context doesn’t determine form; if linguistic form is recognized at all, it must be by largely encapsulated processes” (Fodor 1983: 90). His plea for “encapsulated” subsystems did not rest, in the case of linguistic parsing, on a thorough review of the literature, yet its appeal was considerable, consonant as it was and would be with research conducted in a nativist spirit, for example studies aimed at demonstrating the existence of domain-specific processors, or “proto-theories” at early stages of human development (e.g. Spelke 1988, 1990). The “encapsulation” of syntax as a hypothesis defensible again in psychology should be understood in this wider epistemological context, of which we have been able to offer but a sketch.

## References

- Andresen, Julie Tetel. 1991. Skinner and Chomsky 30 years later. Or: the Return of the repressed. *The Behavior Analyst* 14. 49-60.
- Baars, Bernard J. 1986. *The Cognitive Revolution in Psychology*. New York: The Guilford Press.
- Bach, Emmon W. 1964. *An Introduction to Transformational Grammars*. New York: Holt, Rinehart & Winston.
- Barsalou, Lawrence W. 2008. Grounded cognition. *Annual Review of Psychology* 59. 617-45.
- Bever, Tomas G. 1970. The cognitive basis for linguistic structures. *Cognition and Language Development*, ed. by John R. Hayes. New York: Wiley & Sons. 279-362.
- Bloomfield, Leonard. 1933. *Language*. London: George Allen & Unwin.
- Bloomfield, Leonard. 1936. Language or ideas. *Language* 12(2). 89-95.
- Blumenthal Arthur. 1980 [1970]. *Language and Psychology. Historical Aspects of Psycholinguistics*. New York: Robert E. Krieger.
- Boden, Margaret. 2006. *Mind as Machine. A History of Cognitive Science*. Volume 1. Oxford / New York: Oxford University Press.

- Bolinger, Dwight. 1977. *Meaning and Form*. London & New York: Longman.
- Bransford, John D. & Jeffrey Franks. 1971. The abstraction of linguistic ideas. *Cognitive Psychology* 2. 331-350.
- Bransford, John D. & Marcia K. Johnson. 1973. Considerations on some problems of comprehension. *Visual Information Processing*, ed. by William G. Chase. New York / London: Academic Press. 383-438.
- Breland, Keller & Marian Breland. 1961. The misbehavior of organisms. *American Psychologist* 16(11). 681-4.
- Chase, William G. & Herbert H. Clark. 1972. Mental operations in the comparison of sentences and pictures. *Cognition in Learning and Memory*, ed. by L. Gregg. New York: Wiley. 205-232.
- Chomsky, Noam 1975 [1955]. *The Logical Structure of Linguistic Theory*. New York: Plenum press.
- Chomsky, Noam. 1957a. *Syntactic Structures*. The Hague / Paris: Mouton.
- Chomsky, Noam. 1957b. Review of Charles F. Hockett's *A Manual of Phonology*. *International Journal of American Linguistics* 23(3). 223-234.
- Chomsky, Noam. 1959. Review of B. F. Skinner's *Verbal Behavior*. *Language* 35(1). 26-58.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1972a. *Studies on Semantics in Generative Grammar*. The Hague: Mouton.
- Chomsky, Noam. 1972b. *Language and Mind*. New York: Harcourt Brace Jovanovich.
- Chomsky, Noam. 1980. *Rules and Representations*. New York: Columbia University Press.
- Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin and Use*. Westport, Ct: Praeger.
- Clark, Herbert H. & Eve V. Clark. 1977. *Psychology and Language. An Introduction to Psycholinguistics*. New York: Harcourt Brace Jovanovich.
- Dewey, John. 1896. The reflex arc concept in psychology. *Psychological Review* 3. 357-370.
- Dryer, Matthew S. 1980. The positional tendencies of sentential noun phrases in universal grammar. *Canadian Journal of Linguistics* 25(2). 123-195.
- Fillenbaum, Samuel. 1971. Psycholinguistics. *Annual Review of Psychology* 22. 251-308.
- Fodor, Janet D., Jerry A. Fodor & Merrill F. Garrett. 1975. The psychological unreality of semantic representations. *Linguistic Inquiry* 6. 515-532.
- Fodor, Jerry A. 1965. Could meaning be an  $r_m$ ? *Journal of Verbal Learning and Verbal Behavior* 4. 73-81.
- Fodor Jerry A. 1983. *The Modularity of Mind*. Cambridge, Mass.: MIT Press.

- Fodor, Jerry A., Thomas G. Bever & Merrill F. Garrett. 1974. *The Psychology of Language*. New York: McGraw-Hill.
- Fodor, Jerry A. & Merrill F. Garrett. 1966. *Some Reflections on Competence and Performance*. *Psycholinguistic Papers*, ed. by John Lyons and Roger J. Wales. Edinburgh: Edinburgh University Press. 135-179.
- Fodor, Jerry A. & Merrill F. Garrett. 1967. Some syntactic determinants of sentential complexity. *Perception & Psychophysics* 2(7). 289-296.
- Forster, Kenneth I. & Ilmar Olbrei. 1973. Semantic heuristics and syntactic analysis. *Cognition* 2(3). 319-347.
- Fortis, Jean-Michel. 2015. Generative grammar and cognitive linguistics: On the grounds of a theoretical split in American linguistics. *Metalinguistic Discourses*, ed. by Viviane Arigne & Christiane Migette. Newcastle upon Tyne: Cambridge Scholars Publishing. 53-88.
- Fortis, Jean-Michel. 2018. Prototype theory: On the origins of a theoretical patchwork and its transfer to linguistics. *Theorization and Representations in Linguistics*, ed. by Viviane Arigne & Christiane Migette. Newcastle upon Tyne: Cambridge Scholars Publishing. 168-216.
- Fortis, Jean-Michel & Gilles Col. 2018. Espaces mentaux et intégration conceptuelle: Retour sur la constitution de théories sœurs. *CogniTextes* 18. [<https://journals.openedition.org/cognitextes/1111>, accessed on 03/03/2023. DOI: <https://doi.org/10.4000/cognitextes.1111>].
- Frazier, Lynn. 1988. Grammar and language processing. *Linguistics: The Cambridge Survey*, vol. 2, *Linguistic Theory: Extensions and Applications*, ed. by Frederick Newmeyer. Cambridge, Mass.: Cambridge University Press. 15-34.
- Frazier, Lynn & Charles Clifton Jr. 1996. *Construal*. Cambridge, Mass.: MIT Press.
- Fries, Charles Carpenter. 1952. *The Structure of English. An Introduction to the Construction of English Sentences*. London & Harlow: Longmans Green & Co.
- Givón, Talmy. 1979. *On Understanding Grammar*. New York / San Francisco / London: Academic Press.
- Greenberg, Joseph H. 1966. Some universals of grammar with particular reference to the order of meaningful elements. *Universals of Grammar*, ed. by Joseph H. Greenberg, 2nd edition. Cambridge, Mass.: MIT Press. 73-113.
- Greene, Judith. 1972. *Psycholinguistics: Chomsky and Psychology*. Harmondsworth: Penguin Education.
- Harris, Randy Allen. 1993. *The Linguistics Wars*. Oxford: Oxford University Press.

- Hawkins, John A. 1983. *Word Order Universals*. New York / London: Academic Press.
- Hockett, Charles F. 1953. Review of *The Mathematical Theory of Communication* by Claude L. (sic) Shannon and Warren Weaver. *Language* 29(1). 69-93.
- Holt, Edwin B. 1915. Response and cognition I: The specific-response relation. *Journal of Philosophy, Psychology and Scientific Methods* 12(14). 365-373.
- Honeck, Richard P. 1980. Historical notes on figurative language. *Cognition and Figurative Language*, ed. by Richard P. Honeck & Robert R. Hoffman. Hillsdale, NJ: Lawrence Erlbaum. 25-46.
- Hooper, Joan Bybee & Sandra A. Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4(4). 465-497.
- Hubel David H. & Torsten N. Wiesel. 1963. Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. *Journal of Neurophysiology* 26. 994-1002.
- Hubel David H. & Torsten N. Wiesel. 1965. Comparison of the effects of unilateral and bilateral eye closure on cortical unit responses in kittens. *Journal of Neurophysiology* 28. 1029-1040.
- Huck, Geoffrey & John Goldsmith 1995. *Ideology and Linguistic Theory: Noam Chomsky and the Deep Structure Debates*. New York / London : Routledge.
- Jackendoff, Ray S. 1969. *Some Rules of Semantic Interpretation for English*. Massachusetts Institute of Technology. Dept. of Foreign Literatures and Linguistics. Thesis. Ph.D. [ <https://dspace.mit.edu/handle/1721.1/12965>, accessed 03/03/2023].
- Jacobs, Roderick A. & Peter S. Rosenbaum, ed. 1968. *English Transformational Grammar*. Waltham, Mass.: Blaisdell.
- Jenkins, James J. & David S. Palermo. 1964. Mediation processes and the acquisition of linguistic structure. *The Acquisition of Language: Report of the Fourth Conference Sponsored by the Committee on Intellective Processes Research of the Social Science Research Council, Monographs of the Society for Research in Child Development* 29(1). 141-169.
- Joseph, John E. 2002. How behaviourist was *Verbal Behavior*? *From Whitney to Chomsky. Essays in the History of American Linguistics*. Amsterdam / Philadelphia: John Benjamins. 169-180.
- Joseph, John E. 2018. *Language, Mind and Body. A Conceptual History*. Cambridge: Cambridge University Press.
- Katz, Jerrold J. 1964. Mentalism in linguistics. *Language* 40(2). 124-137.

- Keenan, Edward L. 1973. Logical expressive power and syntactic variation in natural languages. *Formal Semantics of Natural Languages*, ed by Edward L. Keenan. Cambridge: Cambridge University Press. 406-421.
- Keenan, Edward L. 1978. The syntax of subject-final languages. *Syntactic Typology*, ed. by Winfred P. Lehmann. Austin: University of Texas Press. 267-327.
- Keenan, Edward L. 1979. On surface form and logical form. *Studies in the Linguistic Sciences* 8(2). 1-41.
- Kintsch, Walter. 1974. *The Representation of Meaning in Memory*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Klima, Edward S. & Ursula Bellugi. 1966. Syntactic regularities in the speech of children. *Psycholinguistic Papers*, ed. by John Lyons & Roger J. Wales. Edinburgh: Edinburgh University Press. 183-208.
- Knobloch, Clemens. 1988. *Geschichte der Psychologische Sprachauffassung in Deutschland von 1850 bis 1920*. Tübingen: Max Niemeyer Verlag.
- Koch, Sigmund. 1951. Theoretical Psychology 1950: An overview. *Psychological Review* 58. 295-301.
- Koerner, E.F.K. 2004. Linguistics and revolution with particular reference to the Chomskyan revolution. *Language and Revolution / Language and Time*, ed. by Frank Brisard, Sigurd d'Hondt & Tanja Mortelmans. Antwerp: University of Antwerp. 3-62.
- Lakoff, George. 1976 [1963]. Toward generative semantics. *Notes from the Linguistic Underground. Syntax and Semantics* (vol. 7), ed. by James McCawley. New York: Academic Press. 43-63.
- Lakoff, George. 1970. Linguistics and natural logic. *Synthese* 22. 151-271.
- Lakoff, George & Mark Johnson. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.
- Lakoff, George & Mark Johnson. 1999. *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. New York: Basic Books.
- Langacker, Ronald. 1974. Movement rules in functional perspective. *Language* 50(4). 630-64.
- Leahey, Thomas H. 1994. *A History of Modern Psychology*. Englewood Cliffs, NJ: Prentice Hall.
- Lenneberg, Eric H. 1967. *Biological Foundations of Language*. New York: John Wiley & Sons.
- Léon, Jacqueline & Nick Riemer. 2015. Genèse et développement du concept de grammaticalité dans la pensée de Chomsky (1951-1965). *Histoire Epistémologie Langage* 37(2). 115-152.
- Levelt, Willem J.M. 2013. *A History of Psycholinguistics: The Pre-Chomskyan Era*. Oxford: Oxford University Press.

- McCawley, James D. 1968. Lexical insertion in a transformational grammar without deep structure. *Papers from the Fourth Regional Meeting of the Chicago Linguistic Society*. 71-80 (repr. in McCawley 1976. 155-166).
- McCawley, James D. 1971. Prelexical Syntax. *Georgetown University Monograph Series on Languages and Linguistics* 24. 19-33 (repr. in McCawley 1976. 343-56).
- McCawley, James. 1976. *Grammar and Meaning. Papers on Syntactic and Semantic Topics*. New York: Academic Press.
- Miller, George A. 1963 [1951]. *Language and Communication*. New York / Toronto / London: McGraw-Hill.
- Miller, George A. 1962. Some psychological studies of grammar. *American Psychologist* 17(11). 748-762.
- Miller, George A. & Noam Chomsky. 1963. Finitary models of language users. *Handbook of Mathematical Psychology*, vol. 2, ed. by Robert D. Luce, Robert Bush & Eugene Galanter. New York: Wiley. 419-492.
- Miller, George A., Eugene Galanter & Karl H. Pribram. 1960. *Plans and the Structure of Behavior*. New York: Holt, Rinehart & Winston.
- Miller, George & Philip Johnson-Laird. 1976. *Language and Perception*. Cambridge, Mass.: Harvard University Press.
- Miller, George A. & Kathryn McKean. 1964. A chronometric study of some relations between sentences. *The Quarterly Journal of Experimental Psychology* 16(4). 297-308.
- Morton, John. 1969. Interaction of information in word recognition. *Psychological Review* 76(2). 165-178.
- Neisser, Ulric. 1967. *Cognitive Psychology*. New York: Meredith.
- Newmeyer, Frederick. 1986 (2nd ed.). *Linguistic Theory in America*. New York: Academic Press.
- Nielsen, Janet. 2010. *Private Knowledge, Public Tensions: Theory Commitment in Postwar American Linguistics*. PhD Dissertation. University of Toronto.
- Osgood, Charles E. 1952. The nature and measurement of meaning. *Psychological Bulletin* 49(3). 197-237.
- Osgood, Charles E. 1963. On understanding and creating sentences. *American Psychologist* 18(12). 735-751.
- Osgood, Charles E. & Thomas Sebeok, eds. 1965. *Psycholinguistics: A Survey of Theory and Research Problems, with a Survey of Psycholinguistic Research, 1954-1964* by Richard A. Diebold, Jr. Bloomington: Indiana University Press
- Partee, Barbara H. 1979. Semantics - mathematics or psychology? *Semantics from Different Points of View*, ed. by Rainer Bäuerle, Urs Egli & Armin von Stechow. Berlin: Springer-Verlag. 1-14.

- Partee, Barbara H. 2011. Formal semantics: Origins, issues, early impact. *The Baltic Yearbook of Cognition, Logic, and Communication* 6. 1-52. DOI: <https://doi.org/10.4148/biyclc.v6i0.1580>
- Prideaux, Gary D. 1975. Linguistic theories: Where are we going? Paper presented at the Annual Meeting of the Transmountain Regional Conference of the International Reading Association, Calgary.
- Putnam, Hilary. 1960. Minds and machines. *Dimensions of Mind: A Symposium*, ed. by Sidney Hook. New York: New York University Press. 148-179.
- Quillian, M. Ross. 1966. *Semantic Memory*. Cambridge / New York / Chicago / Los Angeles: Bolt Beranek & Newman.
- Radick, Gregory. 2016. The unmaking of a modern synthesis: Noam Chomsky, Charles Hockett, and the politics of behaviorism, 1955-1965. *Isis* 107(1). 49-73.
- Rastier, François. 2010 [1991]. *Sémantique et recherches cognitives*. Paris: Presses Universitaires de France.
- Reber, Arthur S. 1987. The rise and (surprisingly rapid) fall of psycholinguistics. *Synthese* 72. 325-339.
- Riemer, Nick. 2009. Grammaticality as evidence and as prediction in a Galilean linguistics. *Language Sciences* 31. 612-633.
- Roberts, Paul. 1964. *English Syntax: A Programed Introduction to Transformational Grammar*. New York: Harcourt, Brace & World.
- Rosenblueth, Arturo, Norbert Wiener & Julian Bigelow. 1943. Behavior, purpose and teleology. *Philosophy of Science* 10. 18-24.
- Savin, Harris B. & Ellen Perchonok. 1965. Grammatical structure and the immediate recall of English sentences. *Journal of Verbal Learning and Verbal Behavior* 4. 348-353.
- Schank, Roger C. 1972. Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology* 3. 552-631.
- Seuren, Pieter. 1998. *Western Linguistics: An Historical Introduction*. Malden, Mass.: Blackwell.
- Shannon, Claude E. & Warren Weaver. 1949. *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.
- Slobin, Dan I. 1966. Grammatical transformations and sentence comprehension in childhood and adulthood. *Journal of Verbal Learning and Verbal Behavior* 5. 219-227.
- Spelke, Elizabeth S. 1988. The origins of physical knowledge. *Thought without Language*, ed. by Lawrence Weiskrantz. Oxford: Clarendon Press. 168-184.
- Spelke, Elizabeth S. 1990. Principles of object perception. *Cognitive Science* 14. 29-56.

- Tanenhaus, Michael K. 1988. Psycholinguistics: An overview. *Linguistics: The Cambridge Survey*, vol. 3: *Language: Psychological and Biological Aspects*, ed. by Frederick J. Newmeyer. Cambridge: Cambridge University Press. 1-37.
- Thomas, Margaret. 2002. Development of the concept of the poverty of the stimulus. *The Linguistic Review* 19. 51-71.
- Thomas, Owen. 1965. *Transformational Grammar and the Teacher of English*. New York: Holt, Rinehart & Winston.
- Tolman, Edward C. 1932. *Purposive Behavior in Animals and Men*. New York: Century.
- Tomasello, Michael. 2003. *Constructing a Language: A Usage-based Theory of Language Acquisition*. Cambridge, Mass.: Harvard University Press.
- Trabasso, Tom, Howard Rollins & Edward Shaughnessy. 1971. Storage and verification stages in processing concepts. *Cognitive Psychology* 2. 239-289.
- Varela, Francisco J., Evan Thompson & Eleanor Rosch. 1991. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, Mass.: MIT Press.
- Vennemann, Theo. 1974a. Theoretical word order studies: Results and problems. *Papiere zur Linguistik* 7. 5-25.
- Vennemann, Theo. 1974b. Topics, subjects and word order: From SXV to SVX via TVX. *Historical Linguistics 1*, ed. by John M. Anderson & Charles Jones. Amsterdam: North Holland. 339-376.
- Wason, Peter C. 1972. In real life negatives are false. *Logique et Analyse* 15(57/58). 17-38.
- Weiss, Albert Paul. 1929 [1925]. *A Theoretical Basis of Human Behavior*. Columbus, Ohio: R.G Adams & Co.
- Wells, Rulon S. 1947. Immediate constituents. *Language* 23(2). 81-117.

# Complexité et société



# De Bernstein à Labov, la complexité des pratiques langagières : un facteur de hiérarchisation et de discrimination sociales et scolaires

**Marie-Madeleine Bertucci**

CY Cergy Paris Université – EA 7518 LT2D

---

## RÉSUMÉ

L'étude se propose, dans une perspective de compréhension de l'évolution des théories linguistiques, de mettre en évidence les stéréotypes attachés aux parlars non normés et à leur défaut de complexité à partir de la notion de *code restreint* développée par Bernstein (1975) et de celle de *vernaculaire noir américain* élaborée par Labov (1978). Basil Bernstein construit une théorie de l'apprentissage du *langage*, selon laquelle le développement linguistique de l'individu dépend de la classe sociale à laquelle il appartient. Cette analyse lui permet de singulariser la relation au langage de chaque groupe social et d'en déduire l'opposition inégalitaire et clivante fondatrice de sa théorie entre le *langage formel* ou *code élaboré*, apanage des classes sociales les plus favorisées et le *langage commun* ou *code restreint*, parlé par les membres des classes les plus stigmatisées socialement et économiquement. Il construit pour ce faire une typologie des traits linguistiques propres à chacun des deux codes. La théorie du code restreint de Bernstein et les discriminations sociales fortement minorantes qui en découlent ont été vivement critiquées par William Labov, dans les travaux qu'il a consacrés à l'étude de la variété de langue parlée dans les ghettos de New York, dans les années soixante, par les jeunes Noirs. Labov a dénoncé le stéréotype ethnicisé, véhiculé par le discours commun, visant à isoler un parler spécifique identifié comme celui des Noirs, assignés à résidence dans ces lieux urbains ségrégués, qu'il nomme le *vernaculaire noir américain* ou VNA.

Le VNA serait responsable, pour l'opinion commune, du taux élevé d'échec scolaire observable dans ces espaces. Or Labov rejette cette conception et attribue les difficultés scolaires des jeunes de ces ghettos aux différences linguistiques entre leur parler et l'anglais standard des enseignants. Ces différences de structures jointes à des différences de fonctions entre les deux variétés, expliqueraient pour partie les réticences de ces jeunes à apprendre une langue ressentie comme distante, dans un contexte susceptible d'être perçu comme hostile et déstabilisant. Enfin, il dénonce l'illusion que le vernaculaire noir américain constituerait un code autonome distinct de l'anglais standard.

#### **MOTS-CLÉS**

code restreint, code élaboré, vernaculaire noir américain, handicap linguistique, sociolectes, inégalités sociales

#### **ABSTRACT**

The study proposes, in a perspective of understanding the evolution of linguistic theories, to highlight the stereotypes attached to non-standardized languages and their lack of complexity, based on the notion of restricted code developed by Bernstein (1975) and that of Black American vernacular developed by Labov (1978). Basil Bernstein builds a theory of language learning, according to which the linguistic development of the individual depends on the social class to which he belongs. This analysis allows him to singularize the relationship to language of each social group and to deduce the unequal and divisive opposition that is the basis of his theory between formal language or the elaborated code, the prerogative of the most privileged social classes, and common language or the restricted code, spoken by the members of the most socially and economically stigmatized classes. To do this, he constructs a typology of linguistic traits specific to each of the two codes. Bernstein's theory of the restricted code and the strongly minoritizing social discriminations that result from it were strongly criticized by William Labov, in the work he devoted to the study of the variety of language spoken in the ghettos of New York in the 1960s by young Blacks. Labov denounced the ethnicized stereotype, conveyed by the common discourse, aiming to isolate a specific language identified as that of the Blacks, assigned to residence in these segregated urban places, which he called the Black American vernacular. The Black American vernacular would be responsible, according to the common opinion, for the high rate of school failure observable in these spaces. However, Labov rejects this view and attributes the academic difficulties of young people in these ghettos to the linguistic differences between their speech and the standard English of the teachers. These differences in structure, combined with differences in

function between the two varieties, would explain in part the reluctance of these young people to learn a language perceived as distant, in a context likely to be perceived as hostile and destabilizing. Finally, he denounces the illusion that the Black American vernacular constitutes an autonomous code distinct from standard English.

#### KEYWORDS

restricted code, elaborated code, Black American vernacular, language disability, sociolects, social inequalities

---

On se propose de mettre en évidence dans les pages qui suivent l'évolution des théories linguistiques concernant les parlers non normés en partant d'une analyse des différences de perspective de Basil Bernstein (1924-2000) et de William Labov (1927-). On s'efforcera de montrer d'abord en quoi le « code restreint » (1980) de Bernstein et le « vernaculaire noir américain »<sup>1</sup> (1993) de Labov se distinguent en raison d'une problématisation différente de la question des parlers non normés<sup>2</sup>. Ces deux notions servent

---

1. « *Restricted code* » et « *Black American vernacular* », en anglais. Voir aussi infra la note 9 sur cette dénomination.

2. On a fait le choix ici de caractériser ces parlers comme *non normés* de manière à éviter d'utiliser l'adjectif *populaire*, que nous percevons comme fortement polysémique. En effet selon Gadet, qui souligne qu'aucune de ces définitions n'est réellement satisfaisante, *populaire* peut caractériser soit les locuteurs au plan sociologique, soit leurs pratiques langagières et leurs spécificités au plan linguistique (1992 : 24). Le dictionnaire de linguistique de Dubois *et al.* précise à l'article *Populaire* que le terme désigne : « tout trait ou tout système linguistique exclu de l'usage des couches cultivées et qui, sans être grossier ou trivial, se réfère aux particularités du parler utilisé dans les couches modestes de la population » (1973 : 385), ce qui conduit à le définir comme « un usage non standard stigmatisé » déclaré populaire par « le regard social » (Gadet *op. cit.* : 27) qui identifie de la sorte la position sociale modeste des locuteurs (Bourdieu 1983 : 101). Bourdieu voit dans la catégorie populaire une notion « attrape-tout » (*ibid.*) qui contribue à structurer le monde social selon « une taxinomie dualiste » (*ibid.* : 100), opposant la légitimité et la distinction au vulgaire. La notion, d'une manière générale, a des contours flous et constitue un « classificateur déclassant » (Gadet 2003). Nous avons fait le choix de caractériser ces parlers en considérant qu'ils étaient *non normés* car *non normés* paraît plus englobant et moins chargé de connotations péjoratives que *populaire* ou *non standard* si on admet que *non standard* s'oppose symétriquement à *standard*. *Non normé* dans

à évoquer les variétés d'anglais parlées par des jeunes issus de milieux fortement défavorisés socialement et économiquement dans les années soixante en Angleterre et aux États-Unis. La notion de code restreint est utilisée par Bernstein pour désigner le parler des jeunes de la classe ouvrière anglaise de la période concernée. Labov a recours à la notion de vernaculaire noir américain pour dénommer « le dialecte relativement uniforme parlé [...] par la majorité des jeunes Noirs presque partout aux États-Unis » (1993 [1978] : 9) au moment où il décrit cette situation. Il insiste sur le fait que ces jeunes âgés de « huit à dix-neuf ans » vivent dans « les quartiers réservés des grandes villes » et sont « entièrement plongés dans l'univers culturel de la rue » (*ibid.* : 9-10).

En effet, si Bernstein se fonde sur les notions de classe sociale et de culture de classe, Labov quant à lui complexifie la notion de clivages sociaux en la mettant en relation avec celle de clivages ethniques<sup>3</sup>. Dans l'un et l'autre cas, on observe à l'arrière-plan, en deçà de la question linguistique, des préoccupations politiques et sociales renvoyant pour la Grande-Bretagne au feuilletage d'une société hiérarchique compartimentée et rigide<sup>4</sup>, et pour les États-Unis au poids des données raciales et ethniques dans la structuration de la société. On tentera ensuite de voir en quoi cette opposition impacte l'analyse du degré de complexité des variétés non normées proposée par l'une et l'autre conception. On abordera d'abord la théorie de Basil Bernstein en centrant essentiellement l'analyse sur les deux oppositions centrales du code élaboré / langage formel et du code restreint / langage commun<sup>5</sup> qu'on va définir plus précisément dans les lignes qui suivent. On verra dans un second temps la

---

l'usage que nous en faisons ne signifie pas contraire à la norme du standard mais à l'écart de la pression de la norme du standard, non soumis à celle-ci, ce qui n'exclut pas que ces pratiques se conforment à d'autres normes que celles qui régissent le standard. On pense ici à des normes sociales comme celles qui sont partagées par le groupe de pairs.

3. Voir sur ce point Brubaker (2015).

4. Voir sur ce point Hall (2003).

5. « *Elaborated code, formal language et colloquial and non-standard informal language* », en anglais.

perspective de William Labov et l'étude qu'il réalise du vernaculaire noir américain en se fondant sur une analyse dialectologique.

## 1. Déficit de complexité, hiérarchisation des variétés de langue et inégalités sociales et scolaires dans *Langage et classes sociales* de Basil Bernstein

Basil Bernstein construit une théorie sociologique de l'apprentissage<sup>6</sup>, selon laquelle le développement linguistique de l'individu dépend du groupe social auquel il appartient car il considère que « la structure sociale transforme les possibilités linguistiques en codes spécifiques » (1980 : 69) qui s'opposent au plan linguistique par le « degré de probabilité avec lequel on peut prévoir les éléments syntaxiques qui serviront à organiser le discours signifiant » (*ibid.* : 70). Partant du principe qu'un certain nombre de formes du langage dépendent de faits culturels et non individuels, il en déduit que la possession ou la non possession de ces traits langagiers déterminera la possibilité ou l'impossibilité d'acquérir les compétences sociales permettant l'accès à la mobilité sociale. En d'autres termes, la maîtrise du langage assure le succès social et le rapport au langage peut constituer un trait distinctif servant à opposer les différentes catégories sociales. Des usages linguistiques différents apparaissent selon la classe de la société à laquelle le locuteur appartient. À partir de ce constat, il oppose le code élaboré, apanage des milieux favorisés au code restreint, qui caractérise les classes populaires défavorisées et complexifie cette première opposition à partir d'une seconde, celle du langage formel et du langage commun. Le lexique, la syntaxe et « l'expression symbolique des intentions sous une forme verbale » (*ibid.* : 70) sont les composantes essentielles de l'opposition entre les deux codes. Les locuteurs qui utilisent le code élaboré maîtrisent la

---

6. Cette théorie est connue en France notamment par l'ouvrage intitulé *Langage et classes sociales. Codes sociolinguistiques et contrôle social*, publié en 1975 par les Éditions de Minuit dans la collection « Le sens commun » (Bernstein 1980 [1975]). Il regroupe une série d'articles théoriques publiés par Basil Bernstein à partir de 1960 ainsi qu'une partie du premier volume de l'ouvrage *Class, Codes and Control*, qu'il a publié en 1971 à Londres.

complexité de la syntaxe dans toute son étendue, font usage d'un lexique abstrait sans que leur discours soit prévisible à l'avance (*ibid.* : 70). Les locuteurs, qui n'ont que le code restreint à leur disposition, ont une maîtrise réduite de la syntaxe et du lexique, laquelle rend leurs choix discursifs très prévisibles et limite fortement « l'expression symbolique de [leurs] intentions sous une forme verbale » (*ibid.* : 70).

L'un et l'autre code supposent un rapport au langage distinct des locuteurs, que Bernstein analyse à partir de l'opposition du langage formel et du langage commun, dont les corollaires sont respectivement le code élaboré et le code restreint. Le langage formel est « le type de discours dominant dans les classes supérieures » (*ibid.* : 30). Il permet aux locuteurs de « développer une attitude réflexive à l'égard des possibilités structurales d'organisation de la phrase » (*ibid.*). Le langage commun est « le type de discours en usage dans les strates les plus basses de la classe ouvrière » (*ibid.*). Il se caractérise par « la rigidité de [sa] syntaxe », sa prévisibilité et la faible latitude offerte aux locuteurs concernés de procéder à « des choix et des arrangements personnels » (*ibid.* : 70).

## 1. 1. L'opposition code élaboré, langage formel / code restreint, langage commun versus complexité / déficit

### 1. 1. 1. Le langage formel

Le langage formel (1980 : 30) facilite la construction et la hiérarchisation de concepts ordonnés en vue de l'organisation de l'expérience. Sans créer par lui-même des modes de relations complexes, le langage formel permet d'infléchir l'élaboration de la pensée et des sentiments et d'intérioriser la structure sociale du groupe dès que commence l'apprentissage du langage. La précision de l'organisation syntaxique est la caractéristique essentielle du langage formel. Elle permet l'expression de toutes les nuances logiques à travers notamment l'appareil complexe de la subordination. Le langage formel est orienté vers l'expression verbalisée de dispositions personnelles. Il est caractérisé par les traits suivants :

- précision de l'organisation grammaticale et de la syntaxe ;

- nuances logiques et insistance véhiculées par une construction de la phrase grammaticalement complexe, et spécialement par l'utilisation d'une série de conjonctions et de propositions subordonnées ;
- usage fréquent de prépositions qui indiquent des relations logiques, comme de prépositions indiquant la proximité spatiale et temporelle ;
- usage fréquent des pronoms impersonnels « il », « on » ;
- choix rigoureux des adjectifs et des adverbes ;
- impressions individuelles verbalisées par l'intermédiaire de la structure des relations entre les phrases et à l'intérieur de la phrase, c'est-à-dire d'une manière explicite ;
- symbolisme expressif différenciant dans le détail les significations au niveau des phrases, au lieu de renforcer les mots dominants ou d'accompagner les énoncés d'une manière indifférenciée ;
- usage du langage qui rend attentif aux possibilités attachées à un système complexe de concepts hiérarchisés pour l'organisation de l'expérience. (Bernstein 1980 [1975] : 30-31)

### 1. 1. 2. Le langage commun

Le langage commun (1980 : 38 sv.) engage le locuteur vers des concepts descriptifs plutôt qu'analytiques, ce qui limite les champs de l'expérience par rapport auxquels celui-ci apprend à réagir et à se situer. Par ailleurs, le langage commun, pour être appréhendé avec précision, ne peut se dissocier de la relation de communication où il se situe. En effet, il s'agit moins pour le locuteur de développer une démonstration fondée sur une argumentation conceptuelle rigoureuse que d'ancrer dans la relation de communication, la légitimité du discours prononcé. Ceci explique la présence de formules d'appel au consensus, l'utilisation de proverbes, la recherche d'approbation de l'interlocuteur, lesquels constituent autant de substituts argumentatifs.

Ses principaux traits seraient les suivants :

- « phrases courtes, grammaticalement simples, souvent non terminées, à syntaxe pauvre » ;
- « usage simple et répétitif des conjonctions » : « donc », « et » ; des adverbes : « puis », « alors » ;
- faible nombre de propositions subordonnées ;
- « usage rigide et limité des adjectifs et des adverbes » ;

- faible nombre de tournures impersonnelles sur le modèle de « on pourrait penser... » ;
- caractère « catégorique » des énoncés affirmatifs ;
- « rhétorique de l'appel au consensus<sup>7</sup> » par le biais d'énoncés comme : « Tu te rends compte » ; « tu vois » ;
- argumentation et raisonnements à tendance sentencieuse adossés à des proverbes ;
- impressions individuelles demeurant à l'état « implicite » dans la construction de la phrase. (Bernstein 1980 [1975] : 40)

Chacun des deux codes suppose un rapport au langage différent.

## 1. 2. Des modalités distinctes de rapport au langage

Ce type de rapport au langage se fonde sur le degré de proximité ou de distance que le locuteur entretient à l'égard du langage. C'est cette relation qui justifie l'opposition entre le code élaboré et le code restreint. Basil Bernstein comprend l'opposition entre le langage commun et le langage formel comme une opposition socioculturelle. Il pose en principe (*ibid.* : 38) qu'il existe une corrélation entre l'évolution de l'enfant et la structuration de sa vie émotionnelle. Ainsi l'enfant appartenant à un milieu culturel où le rapport au langage est valorisé, va grandir dans un système éducatif rationnel susceptible de lui permettre d'organiser très tôt la totalité de son expérience, en fonction d'un ensemble explicite de buts et de valeurs. La verbalisation des sentiments se voit accorder une place éminente car le mot sert de médiateur entre l'expression du sentiment et les formes socialement reconnues de celui-ci. Le langage autorise le développement des facultés de discrimination. La saisie

---

7. La dénomination « appel au consensus » est discutable. Les marqueurs comme « tu vois ; tu te rends compte » ne sont pas propres au *code restreint*. Très fréquemment utilisés dans le discours oral, on les trouve dans d'autres contextes que ceux que décrit Bernstein. Selon Riegel *et al.* (2018 : 65), ils peuvent être analysés comme des *marqueurs de structuration* des tours de parole à l'oral. Ils relèvent de la dimension phatique de la communication et permettent au locuteur de capter et de conserver l'attention de son interlocuteur. De ce fait, ils sont identifiés comme des *phatèmes*, autrement dit comme des marqueurs de ponctuation des échanges oraux.

de l'expérience s'en trouve facilitée. La compétence à distinguer et à conceptualiser les objets s'affine, ce qui permet de privilégier un système de « relations médiates » (*ibid.*). Dans cette conception, le langage n'est pas un simple outil de communication, il conditionne au contraire le comportement. Très jeune, le locuteur – Bernstein utilise le terme « enfant » - appréhende sa structure sociale et l'intériorise dès qu'il commence à parler. Sensibilisé à la verbalisation, il peut canaliser et différencier ses affects (*ibid.* : 35-36), stabiliser son impulsivité (*ibid.* : 32).

Les groupes d'enfants, n'ayant à leur disposition que le langage commun, appartiendraient à des groupes économiquement défavorisés, comprenant les membres de la classe ouvrière anglaise, certains milieux ruraux et des jeunes marginaux parfois délinquants (*ibid.*).

Bernstein souligne que le langage commun se caractérise par :

- « des attentes [... ou une] anticipation [de l'avenir...] raccourcis » ;
- une « perception du temps » limitée ;
- une instabilité du comportement affectif ;
- une non explicitation du subjectif ;
- une expérience directe de l'affectif qui ne passe pas par la médiation du langage mais par « un symbolisme expressif ». (*ibid.* : 39)

Comme on le voit, les opérations de différenciation assurant la médiatisation de l'expérience – comparaison, inférence, généralisation – ne sont pas prépondérantes. Il en résulte ce que Bernstein nomme « l'impersonnalité » (*ibid.* : 66), laquelle fortifie la forme spécifique de lien social que constitue le groupe (*ibid.* : 67).

### 1. 3. Impersonnalité et identité sociale

L'ensemble de ces caractéristiques contribue à l'instauration d'une certaine impersonnalité. De ce « caractère impersonnel », découle selon l'auteur :

- une forme de comportement social régi par une structure rigide, autoritaire, explicite où les statuts, les rôles, les classes d'âge et les relations coutumières entre ces éléments deviennent des informations déterminantes pour les comportements des agents. (*ibid.* : 47)

L'impersonnalité suscite deux types de réponses qui s'articulent autour d'une forme de solidarité mécanique. Cette solidarité se traduit par la loyauté envers le groupe et par une absence d'engagement individuel au profit de l'investissement dans le groupe et donc un faible exercice de la pensée personnelle (*ibid.* : 66). Si cette situation crée des relations interpersonnelles chaleureuses, elle secrète également une forme de dépendance et de passivité. Le langage commun brime l'expression propre de l'individu parce que celle-ci l'isole du groupe. Il renforce les résistances au changement, le conservatisme qui permettent aux relations sociales communautaires de prévaloir. La communication ne se fait pas alors exclusivement par le langage mais emprunte également le canal de certaines relations sociales et la soumission à la solidarité de groupe. On peut donc dire que le langage commun exprime en permanence les normes du groupe plutôt que l'expérience individuelle de ses membres. Impersonnel, ce langage favorise l'adhésion rigide à des modèles ritualisés (*ibid.*). L'appropriation du langage conditionne la construction et l'appropriation d'une identité sociale. En effet les modalités des rapports sociaux autorisent à opérer un choix quant au « type de code, qui devient alors une expression symbolique de ces rapports et en vient à régler la nature de l'interaction sociale » (*ibid.* : 75). Révélateur de l'origine socioculturelle de l'individu, le processus d'acquisition du langage s'apparente au processus d'acquisition de l'identité sociale : l'intériorisation de la structure sociale devient la « réalité psychologique de l'individu » (*ibid.* : 128). L'échec scolaire assez fréquent, dans les milieux concernés, accentue de plus la faiblesse de la mobilité sociale (*ibid.* : 58) et contribue à renforcer l'homogénéité du groupe (*ibid.* : 50).

### 1. 3. 1. Code restreint et échec scolaire

L'échec scolaire peut s'expliquer par les difficultés rencontrées par les élèves dans l'acquisition du langage et en particulier du processus de conceptualisation, notamment dans les cas où la famille n'attache pas de valeur à ce type de relations. L'aptitude à passer d'un code à l'autre suppose une agilité que ne possèdent pas nécessairement les enfants issus des milieux populaires. La réussite scolaire passant par la maîtrise du langage formel, les enfants inaptes au changement de

code encourrent un risque d'échec plus grand. Sur un autre plan, les difficultés rencontrées tiennent souvent à ce que les contenus abordés n'appartiennent pas à l'environnement immédiat des élèves. Les résistances aux opérations de conceptualisation et de généralisation vont rendre difficiles les opérations de transfert dans des contextes différents et conforter les mécanismes de parcellisation des savoirs. L'absence de continuité entre les attentes de l'enfant et celles de l'école va creuser un fossé, instaurer une distance susceptible de faire naître passivité, indifférence ou opposition violente (*ibid.* : 54-56).

#### 1. 4. Ambiguïtés de l'analyse de Basil Bernstein

Force est de constater que l'auteur se livre à des généralisations excessives lorsqu'il énonce les conséquences sociales des différences langagières qu'il a mises en évidence. D'une part, on peut vite inférer de l'opposition « code restreint »/ « code élaboré », un jugement de valeur stigmatisant et discriminant qui conduit à figer la hiérarchie sociale, ce qui rigidifie les possibilités de mobilité sociale. D'autre part, attribuer à la différence langagière, l'origine des oppositions de classes, situer dans la seule maîtrise du code élaboré, la clé de la mobilité sociale, sont des traits qu'il convient de nuancer. On ne peut universaliser l'idée que tout individu issu d'un milieu populaire ne possède que le code restreint et qu'inversement, tout membre des classes économiquement aisées maîtrise les deux codes et s'avère apte à les utiliser alternativement. Ce serait accorder trop de crédit au déterminisme social. De plus, l'opposition entre l'implicite des impressions individuelles qui parcourt le code restreint et l'ensemble explicite de buts et de valeurs qui structure le code élaboré paraît trop rigide. En effet, il semble difficile d'admettre que tous les locuteurs du code restreint partagent les mêmes difficultés face à l'anticipation de l'avenir, à la perception du temps, à l'explicitation du subjectif, qu'ils aient tous une expérience directe de l'affectif sans passer par la médiation du langage (*ibid.* : 39). Il paraît difficile de suivre Bernstein jusqu'au bout, lorsqu'il conclut à un comportement social rigide, où les statuts et les rôles sont déterminants (*ibid.* : 47), ce qui revient à étiqueter négativement ces groupes en généralisant les notions de *handicap linguistique* et de *privilège culturel* qu'ont

formulées Bourdieu et Passeron (1964) à propos des étudiants<sup>8</sup> et conduit à ignorer nombre de tentatives réussies de mobilité sociale. On suggèrera que l'apport essentiel de Bernstein réside dans la mise en lumière du caractère éventuellement conflictuel du rapport au langage. Le langage apparaît comme un puissant facteur d'intégration ou de division qui joue un rôle déterminant mais non exclusif dans l'itinéraire social de l'individu et dans ses aptitudes à la mobilité.

Les travaux de William Labov permettent de relativiser et de compléter l'approche de Bernstein. Étudiant la langue parlée par les Noirs américains des ghettos urbains, il dénonce (Labov 1993 [1978] : 27 sv.) le stéréotype ethnicisé profondément ancré dans un certain discours commun d'un parler identifié à celui des Noirs qu'il désigne comme le vernaculaire noir américain ou VNA<sup>9</sup>. Le VNA serait responsable, selon le discours commun, du taux élevé d'échec scolaire observable dans ces espaces. Or Labov rejette cette conception et ce pour différentes raisons. Il insiste en premier lieu sur les spécificités du VNA qu'il analyse comme un système qui se distingue des autres variétés d'anglais, au rang desquelles il compte l'anglais de scolarisation, par le nombre de traits langagiers non normés qu'on y observe et qui mettent en évidence son éloignement de l'anglais

---

8. Pour expliquer le faible pourcentage d'étudiants issus des milieux défavorisés à l'université, Pierre Bourdieu et Jean-Claude Passeron ont développé la théorie du *privilège* culturel (1964). Selon eux, la réussite universitaire dépend de la capacité de l'étudiant à s'appropriier et à véhiculer la langue abstraite de l'argumentation. Or, l'appartenance à un milieu aisé, familiarisé de longue date avec le monde universitaire, constitue un atout supplémentaire et un facteur pesant lourdement dans la réussite scolaire et la mobilité sociale. L'ouverture de l'université à des étudiants venus d'horizons diversifiés ne serait qu'un alibi destiné à masquer le processus réel de transmission du pouvoir d'une génération à l'autre. Loin d'être un instrument d'égalité, l'accès aux études serait un élément de discrimination de plus.

9. « Vernaculaire noir américain » est la traduction française de la dénomination labovienne *Black American Vernacular*. Ce sociolecte est également connu sous la dénomination métaphorique et quasiment intraduisible d'*ebonics* – agglutination d'*ebony* et de *phonics* – (Messier 2012 : 2), qu'on pourrait traduire littéralement par « intonations des Noirs », ou sous celle d'*African American Vernacular English* – AAVE – (Poplack 2000) qui signifie littéralement « anglais vernaculaire africain américain ». La dénomination *African American English* (AAE), – anglais africain américain – est également en usage (Green 2002).

standard (1993 [1978] : 71 et sv.)<sup>10</sup>. Il insiste ensuite sur la nécessité de ne pas interpréter ces spécificités du VNA en termes de déficit, d'erreurs et d'infractions à la norme de l'anglais standard (*ibid.*) mais comme les manifestations d'une structure langagière propre à ce dernier non sur la totalité du système mais sur un certain nombre de points de vulnérabilité du « noyau grammatical » de l'anglais standard (*ibid.* : 73). En cela, il s'oppose clairement à Bernstein (Bertucci 1997) comme on va tenter de le montrer dans la deuxième partie.

## 2. La notion de vernaculaire noir américain dans les travaux de William Labov

### 2. 1. De la dialectologie à la sociolinguistique.

#### Affirmation de l'importance du contexte social pour l'étude du parler vernaculaire

En 1972, l'introduction de l'ouvrage intitulé *Language in the Inner City : Studies in the Black English Vernacular* (Labov 1972a ; trad. fr. 1993 [1978] par Alain Kihm) définit le VNA comme « le dialecte<sup>11</sup>

---

10. Cette prise de position a été illustrée par la Controverse d'Oakland, laquelle va être développée infra dans la note 13.

11. La question de savoir si le VNA est un dialecte ou un semi-créole, à savoir un créole stoppé et réintégré partiellement par la langue source (Kihm 1992 : 79), fait l'objet d'un débat nourri (*ibid.*). Alain Kihm souligne que les linguistes Nord-Américains, à partir des années soixante, ont perçu le VNA comme un dialecte, dont la particularité est d'être distinct des dialectes régionaux de Grande-Bretagne. Selon Dillard (1972 : 6, cité par Kihm 1992 : 80), le VNA serait issu d'une variété d'anglais créolisé, lui-même basé sur un pidgin parlé par la population servile et issu de la côte Ouest de l'Afrique (*ibid.*). Labov lui-même a évolué et si dans les années 1972/1978, il penchait pour l'hypothèse du dialecte, à partir des années 1980, il verra dans le VNA les traces d'un créole plus ancien (Labov 1982b : 409, cité par Kihm 1992 : 81). C'est la raison pour laquelle, on rencontre sous la plume de Labov à partir des années 1980, plutôt le terme : *Black English* ou - anglais des Noirs - (Kihm 1992 : 83) que celui de VNA. Le *Black English* désigne spécifiquement le parler des Noirs défavorisés socialement et économiquement (Fromkin, Rodman *et al.* 1997). Kihm rappelle que Mufwene a consacré un article entier en 1992 à une étude de cette situation en affirmant l'existence de trois analyses susceptibles d'expliquer l'origine du VNA. La pre-

relativement uniforme parlé aujourd'hui par la majorité des jeunes Noirs presque partout aux États-Unis » (*ibid.* : 9), en insistant sur le fait que le VNA « définit l'organisation sociale des groupes de pairs [...] en même temps qu'il est défini par elle » (*ibid.*).

En 1982, dans son article « Objectivity and Commitment in Linguistic Science: The Case of the Black English Trial in Ann Arbor », Labov (1982a ; trad. fr. 1982b par Jo Arditty) étoffe la définition précédente en y ajoutant les entrées qui suivent :

1. Le vernaculaire noir américain est un système appartenant à l'anglais et qui possède un ensemble distinct de règles phonologiques et syntaxiques qui convergent désormais sur de nombreux points avec les règles d'autres dialectes.
2. Il a intégré de nombreux traits de la phonologie, de la morphologie et de la syntaxe de l'anglais du Sud ; les Noirs ont en retour exercé une certaine influence sur les dialectes du Sud dans les régions où ils ont vécu.
3. Il porte les traces de la filiation qui le rattache à un créole plus ancien qui se rapprochait davantage des créoles actuels des Caraïbes.
4. Il possède un système aspectuel hautement développé, qui le différencie nettement des autres dialectes de l'anglais, et qui manifeste une évolution continue de sa structure sémantique. (Labov 1982b : 409 cité par Kihm 1992 : 81)

Il rappelle ensuite que ses travaux sur le vernaculaire noir américain prennent leur source dans une recherche commencée en 1965, financée par le Service municipal d'Éducation de la ville de New York (*Office of Education*) et ayant pour objet de mettre en évidence « les différences entre la langue vernaculaire de la partie sud du centre de Harlem et l'anglais standard de l'école » (1993 [1978] : 10). La problématique portait sur le fait de savoir si les *différences dialectales* (*ibid.*) entre les deux variétés étaient la cause (*ibid.* : 10) de l'échec des élèves de la partie sud de Harlem en lecture.

---

mière est que le VNA serait un dialecte anglais qui aurait conservé des traits langagiers qui ne sont plus utilisés par les locuteurs blancs (Kihm 1992 : 142). La seconde analyse fait du VNA un créole issu d'un lointain pidgin à base anglaise parlé en Afrique de l'Ouest dès le xvi<sup>e</sup> siècle (*ibid.* : 143). La troisième analyse considère que le lien avec les langues européennes n'existe qu'en surface et que le VNA est fondé sur un substrat de langues africaines (*ibid.* : 144). Cette dernière hypothèse serait un développement de la seconde analyse qui fait du VNA un créole (*ibid.*).

Dès l'introduction, il souligne qu'une description linguistique dans une perspective dialectologique de « la grammaire du VNA » ne sera pas suffisante. Le VNA a en effet surtout intéressé jusque-là les dialectologues américains (Kihm 1992 : 80). Il convient selon lui de se situer dans une perspective sociolinguistique car la recherche doit aussi porter sur les fonctions du VNA, sur la « manière dont la culture vernaculaire se sert du langage et comment l'habileté verbale s'y développe » (*ibid.* : 11). Il met en évidence la contradiction qui existe entre l'incontestable habileté verbale de ces locuteurs et leur échec en lecture et s'interroge sur les raisons de l'absence de transfert de cette compétence verbale incontestable dans le domaine scolaire (*ibid.* : 11). Cette contradiction lui permet de légitimer l'incursion que la recherche va opérer du côté des « habiletés verbales [...], des actes de parole réalisés en vernaculaire [...], à l'extérieur de l'école aussi bien que de tout cadre institutionnel dominé par les adultes » (*ibid.*). La recherche a pour objet d'étudier le vernaculaire « dans son cadre social » (*ibid.* : 13) de manière à critiquer :

les psychologues scolaires qui défendent la théorie du déficit et considèrent le langage des jeunes Noirs comme inadapté à l'éducation et impropre à la pensée logique [ainsi que] la position de Jensen<sup>12</sup> et d'autres qui prétendent que les enfants noirs sont génétiquement déficients quant à l'aptitude à former des concepts et à résoudre des problèmes. (*ibid.* : 13)

Labov dénonce les stéréotypes attachés au VNA ancrés dans la réalité sociale et le discours commun (*ibid.* : 27 sv.). Ces stéréotypes constituent les fondements sociaux de la perception du VNA

---

12. Arthur Jensen (1923-2012) a été professeur de psychologie de l'éducation à l'université de Californie, Berkeley. On peut lire à son sujet sur le site de l'université les lignes suivantes extraites de sa nécrologie : « *His years at Berkeley substantially defined his theory, which many found to be controversial, that IQ differences can be explained especially by genetic differences among ethnic groups* » pour lesquelles on proposera la traduction suivante : « Il a élaboré à Berkeley l'essentiel de sa théorie du quotient intellectuel, laquelle a été fortement controversée. Selon Arthur Jensen, les différences de QI s'expliqueraient par des différences génétiques entre groupes ethniques ». Labov a fait partie des adversaires de Jensen. L'allusion à Jensen ici est un écho au débat très vif et à la polémique qui existaient à propos de ces questions dans les années 70 aux États-Unis.

dans l'environnement qu'il étudie. L'étude de ce processus de stéréotypage a été développée par Labov dans *Sociolinguistic Patterns* (Labov 1972b ; trad. fr 2019 [1976] par Alain Kihm) où il analyse les usages sociaux de la langue et notamment l'influence des préjugés, des représentations agressives, polémiques, voire racistes sur les représentations et leur force de stigmatisation.

Labov considère que le vernaculaire noir américain, s'il possède un certain degré de systématisme du fait d'une interrelation des règles (1993 [1978] : 73), ne constitue pas un système séparé. Il réfute, et en cela il s'oppose, à l'analyse que développe Bernstein à propos du code restreint, l'illusion que le vernaculaire noir américain constituerait un système linguistique autonome. Il attribue de plus, les difficultés scolaires des enfants des ghettos dont il étudie le parler et l'approche qu'en font les enseignants :

à une situation d'ignorance réciproque<sup>13</sup>, où l'institutrice et l'élève ignorent chacun le système de l'autre, et par conséquent, les règles nécessaires pour traduire un système en son équivalent. (*ibid.* : 28)

---

13. Dans le contexte du District scolaire unifié d'Oakland (*Oakland Unified School District*), situé en Californie, des efforts notoires ont été réalisés pour valoriser l'*ebonics*, comme langue première des élèves Noirs américains et pour adosser à cette langue, antérieure à l'anglais dans le répertoire verbal des élèves, l'enseignement de l'anglais standard à la fois langue 2 de ces élèves et langue de scolarisation. Ce projet a été formalisé dans un manifeste connu sous le nom de « Résolution d'Oakland », laquelle a été adoptée par le Conseil académique de l'établissement le 18 décembre 1996. Ce projet visait à valoriser l'*ebonics* comme medium d'enseignement de l'anglais standard pour atténuer les faits de minoration et de stigmatisation dont il était marqué en tant que variété non normée, de manière à restituer à ce parler toute sa légitimité (Messier 2012 : 1). Ce projet était fondé sur l'idée que les faibles résultats scolaires des élèves et étudiants Noirs résultaient de leur insécurité linguistique et de leur manque de maîtrise de l'anglais de scolarisation (*ibid.* : 3). Des inflexions ont dû être apportées au projet initial dès le 15 janvier 1997 du fait du tollé de protestations qu'il a suscité en particulier dans les médias. La critique essentielle était que le projet, loin de permettre à ces étudiants d'améliorer leurs performances, allait au contraire contribuer à les enfermer dans leurs particularismes langagiers et à les isoler encore davantage, au lieu de les amener à connaître une mobilité sociale (*ibid.* : 4). Le caractère très innovant de cette expérience continue néanmoins à être souligné dans la période contemporaine (*ibid.* : 1).

La résistance des élèves à intégrer les règles de base viendrait des contradictions apparentes entre le système qu'ils utilisent et celui du maître. Ces différences de structures jointes à des différences de statut, – le statut de l'anglais standard n'est pas le même que celui du vernaculaire noir américain –, expliqueraient pour partie les réticences à apprendre une langue éloignée, dans un contexte qui peut être ressenti comme hostile et déstabilisant. Labov considère que l'échec scolaire des locuteurs du VNA est « l'expression du conflit culturel et politique au sein de la classe » (*ibid.* : 13) et que la réussite ou l'échec<sup>14</sup> en lecture sont corrélés à la « plus ou moins grande participation à la culture vernaculaire » (*ibid.*).

## 2. 2. La variation diastratique. Culture des rues et groupe de pairs

Labov rejoint Bernstein par l'importance qu'il accorde aux relations communautaires et à l'influence du groupe de pairs. Il s'attache à analyser les parlers d'enfants appartenant à la « *lower class* »<sup>15</sup> (*ibid.* : 282) où la cellule familiale est organisée autour de la mère en l'absence du père. Il oppose, dans le groupe d'enfants, ceux qu'il estime être des « membres à part entière de l'organisation sociale des rues », intégrés à la culture vernaculaire (*ibid.*) et ceux qui sont détachés de cette culture. Les valeurs du groupe de pairs<sup>16</sup> sont entre autres, la dureté,

---

14. On renverra sur ce point aux conclusions du juge Joiner lors du procès de Ann Arbor, État du Michigan (1979). Celui-ci accordera au VNA un statut de variété non normée d'anglais, suffisamment distincte de l'anglais standard, pour que cette spécificité linguistique puisse peser sur la réussite scolaire des étudiants noirs défavorisés, non censés, selon le juge, connaître aussi celui-ci. Cette décision conduira à la prise en compte par les programmes scolaires de ces particularismes (Kihm 1992 : 86).

15. Classe sociale très défavorisée socialement, économiquement et souffrant d'une discrimination ethnique (notre traduction). Labov donne les indications suivantes à propos de ce groupe social : « Les familles de la *lower class* sont typiquement à base féminine, matrifocales, sans [...] soutien économique régulier » (1993 : 282). Elles vivent dans les « ghettos urbains » (*ibid.*) et sont victimes de ségrégations ethniques, sociales et scolaires (*ibid.*).

16. L'importance du groupe de pairs dans la socialisation des jeunes des milieux défavorisés abordés ici n'est pas propre à leur groupe et ne constitue pas un

la débrouillardise, la recherche de sensations, l'autonomie (*ibid.* : 332). Les aptitudes intellectuelles et le sens de la débrouillardise sont davantage utilisés « en tant qu'instrument pour manipuler autrui, plutôt que comme moyen pour s'informer ou résoudre des problèmes abstraits » (*ibid.* : 331). Dans ces conditions, le savoir de l'école paraît lointain, hostile et étranger et ne peut être une source de prestige contrairement aux attitudes de défi traditionnelles : aptitude à la bagarre, aux insultes rituelles, aux vanes, comportements et actes de parole qui appartiennent à la culture des rues (*ibid.* : 332).

Labov comme Bernstein note qu'appartenir au groupe, c'est :

*adhérer* à cet ensemble de valeurs et *accepter* une somme d'obligations personnelles à l'égard de ceux qui partagent le même environnement et le même système de valeurs. (*ibid.* : 333)

L'appartenance au groupe se manifeste fortement entre 13 et 15 ans, âge qui correspond à la scolarité au collège dans le système éducatif français. C'est dans cette tranche d'âge que l'opposition au langage et les difficultés relationnelles sont souvent les plus aiguës.

Comme on le voit, un certain nombre de points communs réunissent les deux perspectives qui viennent d'être évoquées : valeurs du groupe de pairs et solidarité notamment. Toutefois Labov s'oppose à Bernstein sur deux points : d'une part, il ne traite pas le vernaculaire noir américain comme un système à part entière comme le fait Bernstein avec le code restreint, d'autre part, il n'en

---

type spécifique de socialisation qui serait caractéristique de leurs modalités relationnelles. Le poids du groupe de pairs est aussi important dans la socialisation de jeunes issus de milieux favorisés. On renverra sur ce point, pour les États-Unis, au parler des *valley girls* – littéralement jeunes filles de la vallée –, lesquelles emblématisent un stéréotype de jeune femme américaine légère et folâtre, vivant en Californie dans la vallée de San Fernando à Los Angeles et issue de milieu aisé. Ce parler, aux intonations caractéristiques, est un thème non négligeable de la culture populaire américaine. Il a fait l'objet d'une chanson de Franck Zappa en 1982, et deux films traitent des *valley girls*, celui de Martha Coolidge (1983) et celui de Rachel Goldenberg (2020), remake du premier. On peut penser également à la variété d'anglais connue sous le nom d'*Oxbridge English*, littéralement anglais d'Oxford et de Cambridge, qui au-delà du sociolecte, évoque des spécificités culturelles et civilisationnelles britanniques propres à la socialisation des étudiants dans ces deux universités ainsi qu'à leurs relations interpersonnelles.

tire pas de conclusion généralisante et porteuse de jugements de valeur sur le bon ou le mauvais parler et les aptitudes supposées des locuteurs. L'enjeu est de mettre en évidence la communauté des locuteurs du VNA et d'en faire une communauté linguistique. En effet dans cette perspective, la communauté linguistique est moins l'ensemble des locuteurs parlant de la même manière que ceux qui, malgré la diversité de leurs pratiques, partagent les mêmes normes et les mêmes jugements :

La communauté linguistique se définit moins par un accord explicite quant à l'emploi des éléments du langage que par une participation conjointe à un ensemble de normes. (Labov 2019 [1976]: 187)

### 3. Conclusion

On soulignera pour conclure que les langues ou variétés de langues évoquées ne sont pas sur un pied d'égalité et que les relations hiérarchiques qu'elles entretiennent contribuent à entretenir une relation inégalitaire réelle entre les locuteurs. Les représentations abordées dans ces pages peuvent être interprétées comme particulièrement significatives de cette relation. L'analyse de Bernstein comme celle de Labov, quelles que soient leurs limites, font émerger ces variétés et en en faisant un objet d'étude leur confèrent un statut. Ce statut, même minoré, leur permet de lutter contre le processus d'invisibilisation qui les affecte, notamment par le biais de la contextualisation opérée, qui est un moyen de contrer les dynamiques clivantes qui s'attachent aux catégories sociales mises en altérité, classe ouvrière pour Bernstein, communauté noire américaine précarisée et ségréguée pour Labov.

### Bibliographie

- Bernstein, Basil. 1980 [1975]. *Langage et classes sociales. Codes sociolinguistiques et contrôle social*. Trad. par Jean-Claude Chamboredon. Paris : Minuit.
- Bertucci, Marie-Madeleine. 1997. L'analyse des erreurs : trois théories. *Travaux de didactique du français langue étrangère* 37. 63-88.

- Bourdieu, Pierre & Jean-Claude Passeron. 1964. *Les héritiers. Les étudiants et la culture*. Paris : Minuit.
- Bourdieu, Pierre. 1983. Vous avez dit « populaire » ? *Actes de la recherche en sciences sociales* 46 [L'usage de la parole]. 98-105. [[www.persee.fr/doc/arss\\_0335-5322\\_1983\\_num\\_46\\_1\\_2179](http://www.persee.fr/doc/arss_0335-5322_1983_num_46_1_2179)], consulté le 02/03/2023].
- Brubaker, Rogers. 2015. *Grounds for Difference*. Cambridge, Mass. : Harvard University Press.
- Dillard, Joey Lee. 1972. *Black English*. New York : Random House.
- Dubois, Jean, Mathée Giacomo, Louis Guespin, Christiane Marcellesi, Jean-Baptiste Marcellesi & Jean-Pierre Mével. 1973. *Dictionnaire de linguistique*. Paris : Larousse.
- Fromkin, Victoria, Robert Rodman, Neil Hultin & Harry Logan. 1997. *An Introduction to Language*. Toronto : Harcourt, Brace & Company Canada.
- Hall, Peter. 2003. Le capital social en Grande-Bretagne. *Revue internationale de politique comparée* 10(3). 357-380.
- Gadet, Françoise. 1992. *Le français populaire*. Paris : PUF.
- Gadet, Françoise. 2003. « Français populaire » : un classificateur déclassant ? *Marges linguistiques* 6. 103-115. [[http://www.revue-texto.net/Parutions/Marges/00\\_ml062003.pdf](http://www.revue-texto.net/Parutions/Marges/00_ml062003.pdf)], consultée le 02/03/2023].
- Green, Lisa. 2002. *African American English: A Linguistic Introduction*. Cambridge : Cambridge University Press.
- Kihm, Alain. 1992. Créole ou dialecte : les tribulations du vernaculaire noir américain. *Langages* 108, *Hétérogénéité et variation : Labov, un bilan*, dir. Françoise Gadet. 79-89 [[https://www.persee.fr/doc/lgge\\_0458-726x\\_1992\\_num\\_26\\_108\\_1652](https://www.persee.fr/doc/lgge_0458-726x_1992_num_26_108_1652)], consultée le 02/03/2023].
- Labov, William. 1972a. *Language in the Inner City: Studies in the Black English Vernacular*. Philadelphie : University of Pennsylvania Press.
- Labov, William. 1972b. *Sociolinguistic Patterns*. Philadelphie : University of Pennsylvania Press.
- Labov, William. 2019 [1976]. *Sociolinguistique*. [Traduction de Labov (1972b) par Alain Kihm]. Paris : Minuit.
- Labov, William. 1993 [1978]. *Le Parler ordinaire. La langue dans les ghettos noirs des Etats-Unis*. [Traduction de Labov (1972a) par Alain Kihm.] Paris : Minuit (coll. Le sens commun).
- Labov, William. 1982a. Objectivity and Commitment in Linguistic Science: The Case of the Black English Trial in Ann Arbor. *Language in Society* 11(2). 165 – 201. DOI : <https://doi.org/10.1017/S0047404500009192>

- Labov, William. 1982b. Objectivité et engagement en linguistique : le procès de l'anglais noir à Ann Arbor. Trad. par Jo Arditty. *Grammaire transformationnelle : théorie et méthodologie*, dir Jacqueline Guéron & Thelma Sowley, avec la collaboration de Mirelle Azzoug. Saint-Denis : Université Paris 8. 377-419.
- Messier, Josette. 2012. Ebonics, la résolution d'Oakland et l'utilisation de dialectes non standard en classe. *Les langues anglaises : histoire, diaspora, culture* 3 : 1-10 [<https://jps.library.utoronto.ca/index.php/elhdc/issue/view/1311>, consulté le 01/03/2023].
- Mufwene, Salikoko S. 1992. Ideology and Facts on African American English. *Pragmatics* 2(2). 141-166.
- Poplack, Shana, dir. 2000. *The English History of African American English*. Oxford : Wiley Blackwell.
- Riegel, Martin, Jean-Christophe Pellat & René Rioul. 2018. *Grammaire méthodique du français* (7e éd.). Paris : PUF.

## Sitographie

- Site de l'université de Californie, Berkeley, nécrologie d'Arthur Jensen, [https://senate.universityofcalifornia.edu/\\_files/inmemoriam/html/ArthurR.Jensen.html](https://senate.universityofcalifornia.edu/_files/inmemoriam/html/ArthurR.Jensen.html), consulté le 11/03/2023.

## Discographie

- Zappa, Frank & Moon Unit Zappa. 1982. Valley Girl. In Frank Zappa, *Ship Arriving Too Late to Save a Drowning Witch*. Barking Pumpkin Records, 4 mn 59 s, genre : rock.

## Filmographie

- Coolidge, Martha. 1983. *Valley Girl*. [Metro-Goldwyn-Mayer](#), Atlantic Entertainment Group, 99 mn, genre : comédie romantique.
- Goldenberg, Rachel. 2020. *Valley Girl*. [Orion Classics](#), [Metro-Goldwyn-Mayer](#) et Sneak Preview Productions, 102 mn, genre : remake musical du film *Valley Girl* de Martha Coolidge.



# La simplicité dans les théories syntaxiques et leurs applications pédagogiques dans les années 1930-1980

**John E. Joseph**

Université d'Édimbourg

---

## RÉSUMÉ

L'idée d'une grammaire représente la recherche d'une simplicité essentielle qui sous-tend la complexité langagière. Il n'est pas surprenant alors que les grands pas en avant dans l'histoire de la théorie linguistique ont consisté en des concepts simplificateurs, de la triade conceptuelle des Modistes à la typologie tripartite de Humboldt, puis aux dichotomies structuralistes et à la réduction générativiste des langues du monde à une seule langue humaine. La simplicité n'est pas simple, pourtant. Elle prend diverses formes à diverses époques, et entre les chercheurs d'une même époque. Cet article analyse le rôle joué par la dichotomie simple-complexe dans les théories de Lucien Tesnière (1893-1954), Michael Halliday (1924-2018) et Noam Chomsky, leurs interprétations psychologiques – surtout celle de la théorie hallidayenne par Basil Bernstein (1924-2000) – et les efforts pour les appliquer dans la pédagogie. Dans les trois cas, de bonnes intentions ont été subverties par un concept inadéquat de la simplicité syntaxique. Comment adapter la dichotomie simple-complexe pour atteindre l'égalitarisme éducatif désiré par Tesnière, Halliday et Chomsky ? Faut-il la réimaginer ? Ou l'abandonner entièrement ? Ou ce but louable est-il destiné à rester utopique ?

## MOTS-CLÉS

complexité linguistique, simplicité linguistique, Tesnière (Lucien), Halliday (Michael), Chomsky (Noam), Bernstein (Basil)

**ABSTRACT**

The idea of a grammar entails the search for an essential simplicity underlying the complexity of people speaking. Hence it is not surprising that the great leaps forward in the history of linguistic theory have been in the form of simplifying concepts, from the conceptual triad of the Modistae to the tripartite typology of Humboldt, then on to structuralist dichotomies and the generativist reduction of the world's languages to one single "human" language. Yet simplicity is not so simple: it takes different forms in different periods, and amongst scholars in any given period. This article analyses the role played by the simple-complex dichotomy in the theories of Lucien Tesnière (1893-1954), Michael Halliday (1925-2018) and Noam Chomsky, along with their psychological interpretations – particularly that of Hallidayan theory by Basil Bernstein (1924-2000) – and attempts to apply them in pedagogy. In all three cases, good intentions were subverted by an inadequate conceptualisation of syntactic simplicity. How can the simple-complex dichotomy be adapted so as to achieve the educational egalitarianism desired by Tesnière, Halliday and Chomsky? Must it be reimagined? Or abandoned entirely? Or is this praiseworthy goal destined to remain utopian?

**KEYWORDS**

language complexity, language simplicity, Tesnière (Lucien), Halliday (Michael), Chomsky (Noam), Bernstein (Basil)

---

## 1. Introduction. La nécessité de simplifier

Dans ses deux derniers livres, parus l'année de sa mort, Michel Foucault (1984a et b) a remarqué le retour perpétuel, dans les textes anciens, de règles à répéter chaque jour comme exercices pour l'esprit et le corps, un régime pour le souci de soi. Ce souci n'était pas égoïste, mais fondé sur la certitude que le bonheur de tous dépend du bien-être de chacun ; que pour bien gouverner les autres, il faut d'abord se gouverner ; que pour bien traiter les autres on doit bien se traiter. Quand on enseigne le premier devoir est de montrer aux étudiants comment se soucier d'eux-mêmes. Si la rhétorique est l'art de persuader, cela n'implique pas de faire plier la volonté des autres au détriment de leur propres intérêts, comme le craignaient les socratiques. Utilisée avec soin – reconnaissant que le vrai souci de soi est inséparable du souci des autres – le langage

peut créer et soutenir ce qu'on appelle aujourd'hui (voir Joseph 2018) la « cognition distribuée », avec des « avantages distribués ». Plus les règles de la gestion soigneuse du langage sont simples, et plus leur répétition quotidienne les fait passer de règles mémorisées à des réflexes, incarnés dans le système nerveux, une partie de ce qu'on appelle aujourd'hui la « cognition étendue ».

Il en va de même pour la grammaire, bien que, dans la tradition gréco-latine elle ne soit pas liée à l'éthique de la même façon que l'est la rhétorique. Dans certaines cultures asiatiques, la grammaire avait une fonction éthique profonde : celle d'empêcher la déviation d'une forme fixe prise comme originale et sacrée. Mêmes les analogistes d'Alexandrie, dans leurs disputes contre les anomalistes de Pergame, n'étaient pas aussi attachés à la fixité que les grammairiens sanscrits. À mesure que les dialectes proto-romans se sont écartés progressivement du latin, la grammaire telle que nous la connaissons a pris forme à des fins pédagogiques, et une tension est apparue entre, d'un côté, le besoin et le désir de simplicité pour satisfaire les élèves et leurs professeurs, et de l'autre, le besoin et le désir de complexité pour satisfaire les savants. Cette tension est devenue visible dès lors que la simplicité relative de la grammaire latine de Donat a été suivie par la complexité de celle de Priscien. Ici il s'agit de la complexité de la description grammaticale plutôt que celle de la langue elle-même ; mais avec l'essor de la grammaire spéculative au bas Moyen Âge, où les savants ont commencé à interpréter le langage comme un *speculum*, un miroir de l'esprit, il est devenu moins clair que la complexité de la grammaire puisse caractériser « la grammaire » comme analyse, sans impliquer « la grammaire » comme objet de l'analyse.

La perte des systèmes d'inflexion dans les langues indo-européennes a rendu la syntaxe signifiante, ce qui est en soi une complexification ; et elle a ouvert la voie à la nostalgie d'un passé idéalisé dans lequel la structure du langage aurait été plus pure, plus simple. Le *locus classicus*, ou plutôt *modernus*, de cette nostalgie se trouve chez Wilhelm von Humboldt (1767-1835). Pour lui, le sommet de la perfection linguistique est atteint dans certaines langues mortes-vivantes : le grec et le chinois, non pas tels que parlés par ses contemporains mais tels que préservés dans les manuscrits ; et

surtout le sanscrit, ce zombie de la famille indo-européenne, toujours en cours de production sous une forme traditionnelle, grâce à un investissement culturel profond dans l'éducation d'une élite.

L'idée d'une grammaire représente la recherche d'une simplicité essentielle qui sous-tend la complexité langagière. Il n'est pas surprenant alors que les grands pas en avant dans l'histoire de la théorie linguistique aient consisté en des concepts simplificateurs, depuis la triade conceptuelle des Modistae, jusqu'au XIX<sup>e</sup> siècle et la combinaison des langues dans les familles historiques ; la focalisation sur l'indo-européen depuis Franz Bopp (1791-1867) ; la typologie tripartite de Humboldt ; les lois sans exceptions des néo-grammairiens ; le signe linguistique et les dichotomies de la linguistique saussurienne ; puis les dichotomies structuralistes, le stimulus-réponse des béhavioristes, et le générativisme, qui réduit les langues du monde à une seule langue humaine.

## 2. Tesnière

Humboldt figure dans la première étude de syntaxe de Lucien Tesnière (1893-1954)<sup>1</sup> :

L'épanouissement de la grammaire comparée au XIX<sup>e</sup> siècle a eu pour effet de fournir des bases solides à la phonétique et à la morphologie et de les doter de cadres universellement admis, au moins dans leurs grandes lignes. Il n'en a pas été de même pour la syntaxe. Car si Guillaume de Humboldt a bien **sent** l'existence d'une *innere Sprachform*, personne n'a encore réussi à en préciser la notion ni à en analyser le contenu. (Tesnière 1934a : 219 ; caractères gras dans l'original)

Cette précision est précisément ce que Tesnière entend réaliser. Sa méthode est d'abord d'« isoler les faits syntaxiques par la comparaison directe de langue à langue, de mécanisme à mécanisme », bien qu'on ne le voie pas faire cette sorte de comparaison dans l'article<sup>2</sup> ; puis de « rapprocher ceux qui se ressemblent, à les ranger par affinités, et

---

1. Voir aussi Tesnière (1959 : 13, note 2).

2. Quand il le fait dans Tesnière (1959), Benveniste (1960 : 22) le critique en disant que « L'intervention constante de plusieurs langues à la fois fait qu'on a souvent l'impression d'une syntaxe ou d'une stylistique comparée plutôt que structurale au sens où on l'entend aujourd'hui ».

à découvrir des groupements partiels, puis la clé de voûte qui nous permettra d'apercevoir la structure d'ensemble de l'édifice et de saisir l'économie générale du système » (*ibid.* : 220). Ici le terme économie est à remarquer, étant ambigu entre le sens de l'efficacité, qui suggère la simplicité, et celui d'un système d'échanges réguliers qui peut être très complexe. Mais la recherche d'une « clé de voûte » est sans ambiguïté la recherche d'une analyse simple ; et quand l'économie elle-même devient la clé de voûte de la linguistique d'André Martinet (1908-1999), il spécifie que c'est sa simplification économique du « principe de moindre effort » de Zipf (1949)<sup>3</sup>.

Tesnière commence par ce qu'il va éliminer : les mots-phrases tels que *oui* et *non*, et les « exclamations affectives » (Tesnière 1934a : 221). La raison de leur élimination est qu'ils sont inanalysables, et à quoi sert l'analyse syntaxique sinon à analyser ? Il fait une première division, tacitement saussurienne, entre la syntaxe statique ou catégorique, et la syntaxe dynamique ou fonctionnelle. Puis une contestation particulièrement tesniérienne : « la syntaxe catégorique enseigne surtout à exprimer des idées précises et nuancées », tandis que la syntaxe fonctionnelle « est à la base de l'enseignement pratique et vivant des langues étrangères » (*ibid.* : 223). Malheureusement, dit-il, ce n'est en général pas la façon actuelle d'enseigner les langues ; mais

---

3. Martinet (1952 : 26 ; 2005 [1955] : 271). Selon Peeters (1983 : 105) les premières occurrences chez Martinet du mot économie sont dans Martinet (1938a : 144) et (1938b), où Peeters considère que son sens est assez vague, bien que le deuxième article le lie déjà à « la tendance au moindre effort ». Martinet ne mentionne pas Zipf, dont le livre de 1935 contient plusieurs mentions du « savings of time and effort », notamment dans la section « Principle of Abbreviation ». L'association du moindre effort avec « l'inertie » dans Martinet (1938b) rappelle plutôt Guglielmo Ferrero (1871-1942), qui dans un article de 1894 a proposé une « loi du moindre effort », qui s'applique en combinaison avec « la loi de l'inertie », que Ferrero attribue à son beau-père, Cesare Lombroso (1835-1909, e.g. Lombroso 1890). Dans les études de Lombroso sur la physiognomie criminelle, l'inertie est un facteur important (Lombroso 1876). Ferrero invoque de manière significative l'autorité d'Herbert Spencer, sans nommer une œuvre en particulier. L'usage martinetien classique d'économie commence, selon Peeters, avec Martinet (1952), et ne change pas dans ses ouvrages ultérieurs. Peeters ne mentionne pas la critique cinglante de l'idée martinetienne de l'économie par Grammont (1938 : 210), qui explique peut-être l'absence de cette idée dans les études de Martinet des années 1940.

c'est à la base de sa propre *Petite grammaire russe*, publiée l'année suivante. Le patron de la méthode est astronomique :

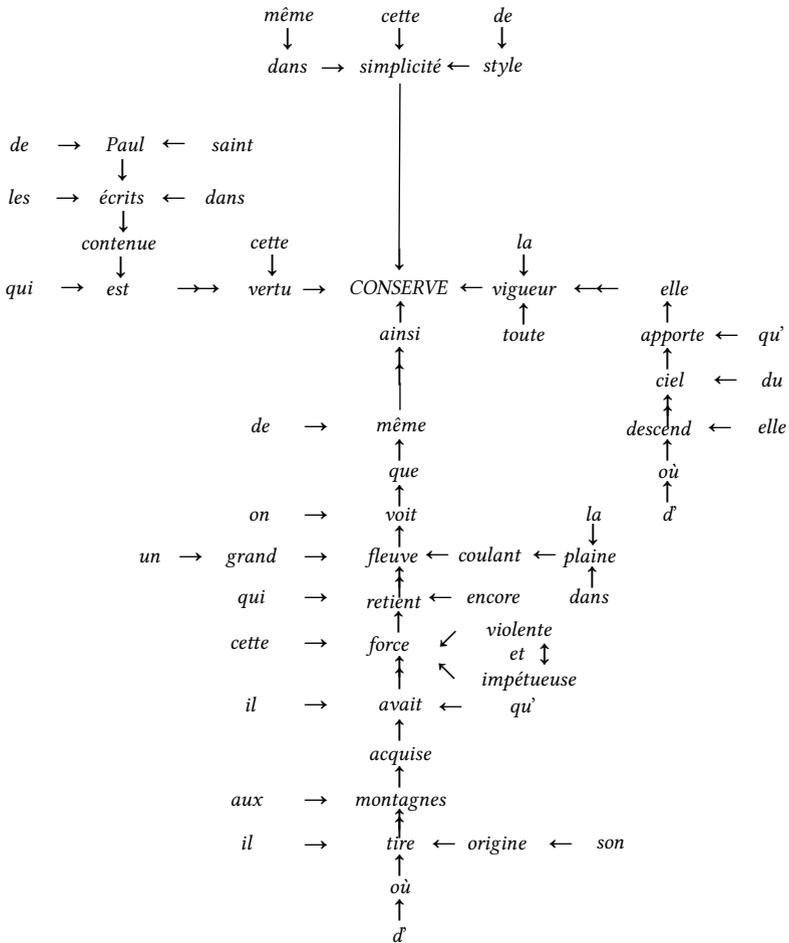
Une phrase se présente comme un système solaire. Au centre, un verbe qui commande tout l'organisme, de même que le soleil est au centre du système solaire. A la périphérie, la foule des éléments grammaticaux, qui sont subordonnés les uns aux autres, et en dernier ressort au centre verbal, selon une hiérarchie à plusieurs étages, tout comme les planètes gravitent autour du soleil et les satellites autour des planètes. Bien entendu, dans cette gravitation universelle de la phrase, il faut, à côté des régissants et des subordonnés de toute sorte, prévoir une place pour les subordonnants, c'est-à-dire pour les éléments qui, n'étant eux-mêmes ni régissants ni subordonnés, ont pour mission de marquer la subordination des autres éléments. Cette réserve faite, toute phrase peut être représentée par un stemma qui indique la hiérarchie de ses connexions. (Tesnière 1934a : 223)

Le « bien entendu » est un tour de force rhétorique : un aveu, déguisé en observation évidente, que la métaphore est trop simple pour capturer la complexité d'une langue ; seul le mot *réserve* en reconnaît la faiblesse, faiblement.

Faire du verbe le cœur de la phrase et même du langage est une démarche traditionnelle. Foucault (1966 : 108) a souligné le fait qu'aux XVII<sup>e</sup> et XVIII<sup>e</sup> siècles le verbe n'est pas traité comme les autres classes de mots : « Le verbe est la condition indispensable à tout discours : et là où il n'existe pas, au moins de façon virtuelle, il n'est pas possible de dire qu'il y a du langage. [...] Adam Smith pense que, sous sa forme primitive, le langage n'était composé que de verbes impersonnels [...], et qu'à partir de ce noyau verbal toutes les autres parties du discours se sont détachées, comme autant de précisions dérivées et secondes ». Le contre-mouvement moderniste aura lieu quand Ogden et Richards prétendront éliminer les verbes du Basic English (voir McElvenny 2018). Cependant, Tesnière (1934a : 224) s'écarte hardiment de la tradition en prenant comme exemple une phrase qui n'est pas simple, mais extraordinairement complexe, du *Panegyrique de saint Paul* de Bossuet, où il s'agit – et peut-être n'est-ce pas un hasard – de la simplicité, avec une dimension céleste :

De même qu'on voit un grand fleuve qui retient encore, coulant dans la plaine, cette force violente et impétueuse qu'il avait acquise aux montagnes d'où il tire son origine : ainsi cette vertu céleste, qui est conte-

nue dans les écrits de saint Paul, même dans cette simplicité de style conserve toute la vigueur qu'elle apporte du ciel, d'où elle descend.



Ce « stemma » n'a qu'un soleil : le verbe principal, *conserve*. C'est le sommet de la simplification, la réduction de cette phrase si complexe à un élément essentiel, un seul mot, auquel tous les autres sont subordonnés.

Dans sa version plus évoluée (Tesnière 1959 : 645), *conserve* sera placé en haut du stemma<sup>4</sup>. Tesnière (1934a : 225) dit qu'« On voit tout l'avantage du procédé, même au point de vue pédagogique, où il permet de souligner que *simplicité* dépend de *conserve* et non de *contenue*, ainsi que pourrait le faire croire une lecture superficielle ». J'accepte ce qu'il dit à propos du fait que *dans cette simplicité de style* modifie *conserve* plutôt que *contenue*, et que le stemma est utile pour montrer cela aux élèves, bien que je doive me borner à accepter la thèse de Tesnière selon laquelle une lecture superficielle suggérerait le contraire ou que l'analyser par stemma révélerait forcément la vraie structure.

Il introduit aussi une forme de transformationnisme, fondé sur le parallélisme qu'il établit entre, d'un côté, les noms et les verbes, et de l'autre, les adjectifs et les adverbes (*ibid.* : 226).

	Action	Substance
Concret.....	Verbe	Substantif
Abstrait.....	Adverbe	Adjectif

Transformer l'adjectif en adverbe par l'addition du suffixe *-ment* est un exercice pédagogique familier. Selon Tesnière,

Il résulte de ce qui précède que l'adverbe est au verbe ce que l'adjectif est au substantif, et que par conséquent, quand on change un substantif en verbe, il faut parallèlement changer l'adjectif en adverbe, ce qu'il est facile de vérifier dans un grand nombre de langues ? [*sic*] Ainsi en français :

*Il comprend clairement.*  
 deviendra *Une compréhension claire. (ibid.)*

Tesnière n'a pas de théorie psycholinguistique de la « nominalisation », comme dans certaines versions de la linguistique chomskyenne, selon laquelle il y aurait un verbe sous-jacent à chaque substantif « dérivé ». Mais le verbe étant le soleil dans le

4. Swiggers (1994 : 215-216, n. 24) trace les différences entre les deux versions dans le contexte du développement de la pensée linguistique de Tesnière, et Mazziotta (2019) analyse le développement de la dimension spatiale des stemmas.

système où s'effectuent ces changements, il devient alors possible de « simplifier » deux catégories en une seule.

Tesnière a introduit l'analyse par stemma dans les cours de préparation d'enseignants qu'il donnait à l'Université de Montpellier. En 1948, un des enseignants, Marcel Cayla, a écrit à Tesnière pour exprimer que, selon lui, le vrai avantage de l'analyse par stemma n'était pas intellectuel, mais sensoriel et physique, donc bien adapté à la « petite intelligence » de ses élèves, qui, quand ils écrivaient, avaient tendance à répéter des phrases lues, comme des perroquets lettrés.

Monsieur Tesnière,

[...] J'ai essayé la méthode des stemma avec mes élèves de français. Parmi les bons élèves elle est acceptée facilement à part quelques rares apprentis qui n'y ont pas réussi du premier coup et ont trouvé les raisins trop verts. Avec les élèves moins doués, ceux qui ont échoué au certificat d'études primaires, cette méthode me paraît offrir de sérieux avantages. Elle pourrait supprimer le psittacisme, si néfaste chez eux. Après avoir vu quel méli-mélo me présentaient ces pauvres apprentis dans leurs analyses grammaticales, j'ai été agréablement surpris des résultats obtenus grâce aux stemmas. Les élèves reprennent le contrôle de leur petite intelligence et, au lieu d'aligner des mots sans suite – et d'ailleurs incompréhensibles pour eux souvent – ils « dessinent » leurs phrases avec entrain. Quelle chance pour eux qui n'aiment pas écrire de pouvoir utiliser en français comme à l'atelier le principal mode d'expression de l'ouvrier : le schéma. Quant au maître, il peut expliquer sans grands développements la phrase qui arrête ses élèves dans une lecture : quelques traits au tableau et le sens apparaît parce qu'on a découvert les connexions<sup>5</sup>.

Tesnière a répondu avec sympathie, offrant toute l'aide qu'il pouvait apporter ; je peux presque lui pardonner de ne pas s'être distancié des remarques de Cayla, en disant par exemple – c'est moi qui imagine cette réponse – « Monsieur, je me permets de vous envoyer un livre de Piaget, qui approuverait ce que vous dites à propos des avantages des schémas, tout en refusant de mesurer l'intelligence des élèves par leur production verbale, et la caractérisant en termes

---

5. Marcel Cayla à Tesnière, Lamalou (Hérault) le 15 mars 1948, dans BnF NAF 28026 (49), Fonds Lucien Tesnière, Correspondance.

non quantitatifs mais temporels, reconnaissant que les individus n'arrivent pas au seuil de chaque stade de développement au même âge, mais finalement ils y arrivent tous ».

### 3. Bernstein

Néanmoins, mêmes les discours « progressistes » sur la langue et l'éducation n'étaient pas à l'abri d'associer la complexité verbale avec une vaste gamme de traits sociaux et cognitifs désirables, et la simplicité avec le déficit. Qui, après tout, aurait pu être plus progressiste que le linguiste marxiste M. A. K. Halliday (1925-2018)<sup>6</sup> et son camarade sociologue de l'éducation, Basil Bernstein (1924-2000), avec sa formule célèbre des codes élaborés et restreints (*elaborated and restricted codes*)? Selon Bernstein (1964), les enfants des couches sociales moyennes et supérieures habitent un milieu où se parle le code élaboré, ce qui leur confère un avantage lorsqu'ils arrivent à l'école. Leur succès, et l'échec de leurs camarades de la classe ouvrière, est rendu presque inévitable par le fait que seul le code élaboré permet le développement d'une vraie identité personnelle et d'une cognition complète de leur environnement. Le code restreint, par contre, produit une forte identité sociale, partagée avec les autres enfants qui ne parlent que le code restreint ; mais c'est la pensée originale, donc individuelle, qu'on valorise dans l'éducation. Le code restreint se réalise dans le contexte d'« une identité culturelle qui réduit le besoin chez les locuteurs d'élaborer leurs intentions verbalement et de les expliciter » ; « sa fonction est de permettre la signalisation d'une identité sociale plutôt qu'individuelle. [...] Le code renforce la solidarité avec le groupe en limitant la signalisation verbale de la différence personnelle. [...] Un fort sentiment d'identité sociale est induit probablement au prix d'un sentiment d'identité personnelle » (Bernstein 1964 : 58, 63).

Dans la présentation initiale de cette idée, où il s'appuie fortement sur l'autorité de Benjamin Lee Whorf (1897-1941) en affirmant le pouvoir du code pour former la manière d'apercevoir et de penser, Bernstein utilise le terme « langage public » pour ce que, six ans plus tard, il appellera le « code restreint ». Tandis que « La parole

---

6. Sur le marxisme de Halliday voir Christie (2007 : 3).

de la classe moyenne est riche en qualifications personnelles et individuelles, et sa forme implique des ensembles d'opérations logiques avancées » (Bernstein 1958 : 164 ; les traductions des textes anglais sont les miennes), le langage public présente d'autres caractéristiques :

phrases courtes, grammaticalement simples, souvent inachevées avec une construction syntaxique médiocre ; emploi simple et répétitif des conjonctions (*so, then, and*), par conséquent les modifications, les qualifications et l'importance logique auront tendance à être indiquées par des moyens non verbaux ; emploi fréquent d'impératifs et de questions courtes ; emploi rigide et limité d'adjectifs et d'adverbes ; emploi peu fréquent du pronom impersonnel (*it, one*) comme sujet d'une phrase conditionnelle ; [...] La sélection individuelle au sein d'un groupe de tours traditionnels joue un grand rôle. Le symbolisme est d'un faible ordre de généralité. *La qualification personnelle est exclue de la structure de la phrase, c'est donc un langage à signification implicite.* (*ibid.* : 164, n., italiques dans l'original)

Phrases grammaticalement *simples*, emploi *simple* des conjonctions – exactement ce que les manuels conseillent aux écrivains apprentis. La dernière phrase, que Bernstein considérait évidemment l'échec et mat, est, autant que je puisse voir, un *non sequitur* : car pourquoi une phrase aurait-elle une signification explicite si sa structure comprend une qualification personnelle, et dans le cas contraire une signification implicite ? Bernstein parle de « langage public » (*public language*) puisqu'il y manque les dimensions langagières qui permettent aux locuteurs de construire un « soi » dans leur parole. Il croit que ce manque limite la capacité de l'enfant de la classe ouvrière à percevoir le monde :

Puisque la nature de leur langage limite les communications verbales des sentiments, ceux-ci ont tendance à être aussi indifférenciés que le langage. Par conséquent, la différenciation émotionnelle et cognitive de l'enfant de la classe ouvrière est relativement moins développée, et les signaux auxquels il répond dans l'environnement seront principalement d'un ordre qualitativement différent. (*ibid.* : 168)

Dans un article de 1959, Bernstein passe en revue les caractéristiques du langage public énumérées ci-dessus et explique comment chacune limite la perception et la pensée des enfants de la classe ouvrière qui le parlent. Bernstein commente positivement

le fait que le langage public « contient sa propre esthétique, une simplicité et une franchise d'expression, émotionnellement virile, concise et puissante et une gamme métaphorique d'une force et d'une pertinence considérables » (Bernstein 1959 : 322-323), et il déclare que le but de l'éducation ne devrait pas être de le remplacer, mais de le préserver tout en permettant aux individus d'utiliser également le « langage formel » de la classe moyenne. Le fait qu'il ne préconise pas la scolarisation des enfants de la classe moyenne en « langage public » fait cependant sonner creux ces éloges, qui sont de plus condescendants<sup>7</sup>. Le passage terminologique de public-formel à restreint-élaboré montre clairement que le contraste entre la complexité et la simplicité est au cœur de cette distinction, et que la simplicité équivaut au déficit.

À son crédit, Bernstein a ensuite infirmé ses « tentatives de spécifier les caractéristiques linguistiques pertinentes associées aux orientations de codage contrastées », après « s'être rendu compte qu'elles étaient erronées » (Williams 2015 : 336). Cela n'élimine cependant pas leur intérêt historique : pourquoi semblaient-elles si vraisemblables à l'époque, non seulement pour Bernstein mais pour la plupart de son vaste public ? Seules les caractéristiques grammaticales spécifiques ont été infirmées par Bernstein, non le concept global de codes élaboré et restreint, qui continue de sous-tendre l'analyse du discours critique (ADC) qui s'est développée directement à partir de l'enseignement de Halliday. La classe dirigeante, c'est-à-dire les gouvernements et les sociétés commerciales, utilise le langage et toute la gamme des ressources multimodales pour opprimer la classe ouvrière, qui est incapable de reconnaître cette oppression linguistique et de résister à ses effets sur sa pensée. Cela explique pourquoi la classe ouvrière continue d'élire des gouvernements qui les opprimeront. La mission de l'ADC est d'exposer les mécanismes de cette oppression, et ainsi de contribuer à défaire son pouvoir.

Si les analystes du discours peuvent accomplir cette mission, c'est parce qu'ils ont bénéficié d'une éducation d'élite avancée, avec la maîtrise du code élaboré, leur permettant de reconnaître la supercherie. Dans l'imagination de ces analystes, les gens de la rue

---

7. Et quelque peu mystérieux : que veut dire « émotionnellement viriles », étant donnée l'association de la virilité avec l'absence d'émotion ?

ne peuvent pas le faire : leur code restreint en fait les marionnettes des maîtres qui font appel à leurs émotions plutôt qu'à leur intelligence, quelles que soient ses « limites ». Ils sont effectivement hypnotisés par des slogans, dans une « tyrannie du simple », pour citer un slogan des opposants au populisme<sup>8</sup>. Si seulement ils étaient capables de lire des milliers de pages de tracts politiques et de manifestes, dans toute leur complexité, ils verraient la lumière... L'analyste du discours contrôle non seulement le code élaboré, mais surtout la grammaire systémique-fonctionnelle de Halliday, qui offre la clé pour déverrouiller le fonctionnement du code. Nombreux sont certainement les lecteurs de ces lignes à partager cette perspective, surtout en ce qui concerne le pouvoir du slogan « simple ». J'indiquerai simplement notre tendance à remarquer ce pouvoir dans le cas du slogan d'un parti gagnant ou d'un produit réussi, et à ignorer les cas beaucoup plus fréquents de slogans de partis perdants ou de produits qui ne se vendent pas bien. Nous devrions nous demander si c'est le slogan qui réussit à cause de ce à quoi il se réfère, où l'inverse ; et si, quand nous dédaignons la sagesse de la multitude, comme Cayla dédaigne l'intelligence de ses élèves moins doués, ce n'est pas nous qui montrons une intelligence faible.

#### 4. Chomsky

En France, beaucoup plus que dans le monde anglophone, l'analyse du discours est associée à Zellig Harris (1909-1992), qui a créé le terme. Son étudiant Chomsky – qui prétend que dans les cours de linguistique de Harris, tout ce qui a été fait est l'analyse de tracts politiques (voir Barsky 2011) – a poursuivi depuis les années 1960 l'analyse politique des textes dans l'esprit de Harris, sans lien avec l'ADC, ni même avec sa propre linguistique générative, insiste-t-il.

---

8. L'internet contient des dizaines de citations d'une pseudo-citation de Jakob Burckhardt (ou d'Edmund Burke) selon laquelle « L'essence de la tyrannie est le déni de la complexité » ou « Le déni de la complexité est le début de la tyrannie ». Aucune source n'en est jamais spécifiée, ce qui rend l'attribution douteuse, mais cela n'en diminue pas sa force pour ceux qui l'évoquent en opposition à une supposée mentalité de troupeau, ou bien un consensus démocratique, qu'ils ne partagent pas. Dans un contexte scientifique, Ball (2016) discute « la tyrannie des explications simples », en se concentrant sur le célèbre « rasoir d'Ockham ».

Joseph (2004) montre comment, au contraire, elles sont étroitement liées, par la théorie chomskyenne de l'interprétation linguistique plutôt que par celle de la production.

Bien des linguistes ont discuté les points de contact entre la grammaire transformationnelle-générative et l'analyse syntaxique de Tesnière<sup>9</sup>. L'une des choses qui les lient est la présomption qu'au-dessous de la complexité de la syntaxe il y a une simplicité latente. Chomsky considère une phrase passive telle que *Bill was hit by John* comme complexe ; le maximum de simplicité se voit dans la phrase active, *John hit Bill*. Dans l'esprit de celui qui produit une phrase passive et de ceux qui la comprennent, tout commence et finit avec son équivalent actif. J'ai déjà mentionné que la « théorie standard » de Chomsky comprenait une transformation dite « nominalisation », par laquelle un nom est dérivé d'un verbe, comme dans l'exemple fourni par Tesnière, *une compréhension claire* et *il comprend clairement*. Cette transformation avait été formulée dans la première thèse de doctorat dirigée par Chomsky (Lees 1960). Chomsky écrit :

Lorsque les processus de dérivation sont productifs, ils ne posent en fait aucune difficulté sérieuse. Considérons, par exemple, des transformations de nominalisation du type de celles qui forment les phrases « leur destruction du bien ... », « leur refus de participer ... », etc. De toute évidence, les mots *destruction*, *refus*, etc., ne seront pas inscrits comme tels dans le lexique. Au lieu de cela, *détruire* et *refuser* seront inscrits dans le lexique avec une spécification de fonctionnalité qui détermine la forme phonétique qu'ils prendront (par des règles phonologiques ultérieures) lorsqu'ils apparaîtront dans des phrases nominalisées. (Chomsky 1965 : 184, ma traduction)

Par « productifs » dans la première phrase, il veut dire productifs de forme régulière, de telle sorte qu'une spécification simple rendra compte de tous les mots dérivés, ou plutôt d'un assez grand nombre d'entre eux pour convaincre que les autres soient marginalisés comme des exceptions. Au fur et à mesure que la discussion se poursuit, cependant, nous voyons des indices que Chomsky est conscient de la façon dont cette productivité est limitée par les processus de dérivation, et malgré le « de toute évidence » de la déclaration précédente, son hésitation et son ambivalence sont palpables. Elles

---

9. Pour des exemples voir Arrivé (1969 : 38-39).

deviendront doctrinales lorsque Chomsky (1970 [1968] : 215) changera d'idée et restreindra la nominalisation aux « nominaux géronatifs » tels que *his comprehending*, l'excluant des « nominaux dérivés » tels que *his comprehension*. Le même article indique clairement que de telles décisions équivalent à un compromis entre simplicité et complexité entre les différentes parties du système :

En général, on peut s'attendre à ce que l'enrichissement d'une composante de la grammaire permette une simplification dans d'autres parties. Ainsi, certains problèmes descriptifs peuvent être résolus en enrichissant le lexique et en simplifiant la composante catégorielle de la base, ou inversement ; ou en simplifiant la base au prix d'une plus grande complexité des transformations, ou inversement. (*ibid.* : 185)

Lees avait donné à Chomsky (1957) son coup de pouce initial avec un compte rendu dans *Language* (Lees 1957) ; ce sera ensuite au tour de Newmeyer (1971), étudiant de Lees, de défendre dans la même revue la principale contribution de son ancien professeur contre la critique de Chomsky. À cette période, les « guerres de la linguistique » (Voir Harris 2021) étaient à leur comble, et l'Université de l'Illinois, où Newmeyer a fait ses études sous la direction de Lees, était l'un des centres de la sémantique générative, beaucoup moins préoccupée de la régularité des règles transformationnelles que l'orthodoxie du MIT. En effet, c'est le travail en sémantique générative, y compris celui de Lakoff sur la métaphore, qui chevauche la première ADC d'une manière jamais égalée par la syntaxe chomskyenne ; et la nominalisation devient une préoccupation de l'ADC, qui la trouve suspecte en raison de sa capacité à camoufler le sujet, quand par exemple *le harcèlement des jeunes noirs* masque le rôle de la police qui les harcèle. Les analystes du discours utilisent le terme « nominalisation » sans prétendre qu'à un certain niveau profond le nom *harcèlement* soit le verbe *harceler*<sup>10</sup> ; cependant, les pistes sont brouillées par le fait qu'ils visent à restaurer des élisions, et qu'il est important pour eux de formuler ces élisions en termes grammaticaux – car ils s'inquiètent, non sans raison, du potentiel reproche que ce

---

10. Ainsi van Dijk (2008 : 827) : « This does not mean that an 'underlying' verb is transformed by the language user, but that a nominalization is chosen directly from the lexicon as the obligatory, the shortest or the easiest way to describe the action ».

qu'ils font n'est pas de la linguistique scientifique, mais une critique textuelle politique.

Réduire deux catégories fondamentales comme substantif et verbe à une seule, avec une opération de transformation, constitue une simplification surtout dans une entreprise qui valorise, non pas l'adéquation descriptive, mais l'adéquation explicative (Chomsky 1965). La réduction explique, ou paraît expliquer, la relation entre substantif et verbe, tout en simplifiant la structure de la « grammaire universelle ».

## 5. Applications théoriques et méthodologiques de la simplicité

Au cours des dix années entre 1954 et 1964, Tesnière élabore la version finale de sa théorie, publiée à titre posthume ; Bernstein formule sa distinction entre langages public et formel, et l'affine en codes restreint et élaboré, en collaboration avec Halliday ; et Chomsky développe sa grammaire transformationnelle-générative jusqu'à aboutir à une « théorie standard ». Chacune de ces entreprises avait à sa base une conception de la simplicité syntaxique avec des conséquences pédagogiques directes – bien que Chomsky fasse bande à part, n'ayant pas poursuivi ces conséquences dans son propre travail ; beaucoup d'autres étaient prêts à le faire. Chomsky insiste sur le fait que sa propre éducation était une perte de temps, en particulier en ce qui concerne la langue, car il était devenu, comme tout enfant, un locuteur à part entière de sa langue maternelle, grâce à un processus biologique achevé vers l'âge de quatre ans. Cela le met en complet désaccord avec Bernstein et Halliday, pour lesquels seul le code élaboré est une vraie langue, ce dont un individu a besoin pour réaliser son potentiel à être un vrai sujet ; le code restreint limite ses locuteurs au statut d'objet collectif. Pour Chomsky, s'inspirant de Humboldt, chaque enfant possède une créativité linguistique infinie, malgré le fait qu'il n'entende parler qu'un code restreint autour de lui – la « pauvreté de l'expérience » de Chomsky<sup>11</sup>, qui prouve la nature biologique du langage et expose ce qui est à son avis l'erreur

---

11. Chomsky (1971 : 17), « poverty of experience », reformulé par la suite en « poverty of the stimulus ».

fondamentale de Bernstein et Halliday : ne pas faire la distinction entre performance et compétence.

La différence coïncide avec leurs opinions politiques : Chomsky l'anarchiste, qui se méfie des institutions publiques, y compris les écoles, contre Halliday le marxiste, qui croient que seul l'État centralisé sera capable de sauver la classe ouvrière de son oppresseur bourgeois, une fois que la révolution aura pris son contrôle de l'État. L'une et l'autre parties comprennent le langage en termes de simplicité et de complexité, mais de manières conceptuellement différentes, dont ils tirent des conclusions diamétralement opposées.

Halliday n'a jamais renié les idées de Bernstein, lesquelles ont leur héritage au sein de l'ADC<sup>12</sup>. Sa grammaire systémique-fonctionnelle a un objectif réducteur et simplificateur qu'elle partage avec les grammaires et l'analyse linguistique en général ; il prétend rejeter le déterminisme structuraliste en faveur des choix faits par les utilisateurs de la langue, mais la revendication dépend en partie d'une exagération du déterminisme structuraliste, et en partie de « choix » chimériques. Je voudrais à présent jeter un coup d'œil à un passage assez révélateur de son livre de 1975, *Learning How To Mean* (Apprendre à signifier), la transcription d'un échange verbal entre Halliday et Ruqaiya Hasan et leur fils :

**Father:** you went on a train yesterday

[*tu as pris le train hier*]

**Nigel:** too-too, bah-bah (= « when I got off, the train went away and I waved to it »)

**Father:** and you said bye-bye to the train

[*et tu as dit bye-bye au train*]

**Nigel:** ahn (= « another »)

**Father:** and you saw another train?

[*et tu as vu un autre train ?*]

**Nigel:** (long list of things seen then:) wla (« flag »)

**Mother:** oh, you saw some flags?

[*tu as vu quelques drapeaux ?*]

**Nigel:** (holding out palm) gra (« gravel »)

---

12. Sur les origines hallidayennes de l'ADC, voir par ex. Blommaert & Bulcaen (2000 : 454).

**Mother:** and you had some gravel

[*et tu avais du gravier*]

**Nigel:** (touching palm, lips rounded, very quiet) ooh

**Father:** and you hurt your hand with the gravel?

[*et tu t'es fait mal à la main avec le gravier ?*]

**Mother:** no, that was with the stick, the one with prickles on

[*non, ça c'était avec le bâton, celui avec les piquants*]

**Nigel:** blah (« blood »).

**Mother:** and there was blood on it, yes

[*et il y avait du sang dessus, oui*]

(Halliday 1975 : 90)

Les énoncés de Nigel ne pourraient guère être plus simples : neuf syllabes, étendues jusqu'à 52 dans l'interprétation de Father et Mother. Halliday dit qu'à ce stade initial il n'y a pas de « niveau de forme linguistique ».

C'est un système fonctionnel, dans lequel la fonction est identique à l'utilisation ; chaque énoncé n'a qu'une seule fonction, et les significations sont de l'ordre de « donne-moi ça, je suis intéressé, soyons ensemble ». Les fonctions initiales sont instrumentales, réglementaires, interactionnelles et personnelles ; suivent ensuite l'heuristique et l'imaginaire. Chaque élément de la langue est une simple paire contenu-expression ; il n'y a pas de niveau de forme linguistique (pas de grammaire).

À un certain stade, l'enfant commence à utiliser le langage dans une fonction « mathétique » à des fins d'apprentissage. Cela se présente comme une généralisation du personnel et de l'heuristique ; le langage dans l'identification du soi et, comme corollaire, dans l'exploration du non-soi. (Halliday 1975 [2004] : 52, ma traduction)

Cette langue initiale, la simple paire contenu-expression avec une seule fonction, se complexifie lorsque le langage est utilisé « dans l'identification du soi », et c'est ici que les codes élaboré et restreint de Bernstein commencent à être ressentis comme problématiques, si effectivement le code restreint est un obstacle à l'identité personnelle. Ce n'est évidemment pas un problème pour Nigel. Je suppose que l'expansion de ses énoncés par Father et Mother implique leur présomption soit que la pensée est entièrement formée dans l'esprit de Nigel, et qu'il n'est pas encore capable de l'exprimer ; ou qu'elle n'est pas entièrement formée, juste sous une forme rudimentaire,

peut-être une image visuelle, dont il ne connaît les mots que pour quelques éléments. Il lui suffit de leur faire comprendre ce qu'il veut dire – de toute façon, ils croient comprendre. Eh bien, ce sont ses parents – ils interagissent avec lui depuis la naissance, même avant, dans le ventre – et personne n'est plus qualifié qu'eux pour interpréter *too-too bah-bah* ou *gra*. À quoi sert la traduction en anglais adulte de chaque énoncé ? Est-ce pour Nigel, pour son processus d'apprentissage ? Pour l'autre parent ? Pour le magnétophone et les recherches de Halliday ?

Quelle que soit la réponse, cet « apprentissage de la signification » semble tout à fait favorable et bienfaisant à travers le prisme hallidayen. Il va de pair avec l'étayage du vieux camarade d'armes (contre B. F. Skinner) de Chomsky, Jerome Bruner (1915-2016), qui avait cependant rompu avec Chomsky sur la question de savoir si le dispositif d'acquisition du langage de Chomsky, le LAD (*Language Acquisition Device*), nécessitait ce que Bruner a appelé un système de support d'acquisition du langage, la LASS (*Language Acquisition Support System*<sup>13</sup>) – ou si une LASS, qui est directement observable, devait supporter les dépenses d'un LAD hypothétique. Mais à travers un prisme chomskyen, « apprendre à signifier » a un caractère sinistre : car si les interventions de Father et Mother contribuent à l'apprentissage de la signification par Nigel, où est alors la frontière entre cela et la « fabrication du consentement » que Chomsky attribue à la propagande des gouvernements et des sociétés commerciales ?

Selon Chomsky, un enfant a effectivement besoin d'« input », l'entrée des données, pour acquérir les éléments lexicaux et syntaxiques qui fixeront les « paramètres » et lui permettront de parler et de comprendre la langue particulière de son environnement ; mais c'est un processus automatique, insensible à tout contexte de développement, en dehors des anomalies extrêmes. Ce point de vue a été didactisé pour des générations d'étudiants débutants en linguistique en Amérique du Nord à travers le manuel de Fromkin *et*

---

13. Il s'agit d'un jeu de mots de la part de Bruner : *Ilka lassie has her laddie* (Chaque fille a son garçon), du poème *Comin' Thro' the Rye* (1782) de Robert Burns, barde national de l'Écosse.

*al.* (2014 : 424)<sup>14</sup>, largement utilisé et axé sur la grammaire générative, qui comprend cet extrait de dialogue mémorable :

CHILD: Nobody don't like me.

[*Personne m'aime pas.*]

MOTHER: No, say « Nobody likes me. »

[*Non, dis « Personne ne m'aime ».*]

CHILD: Nobody don't like me.

(dialogue repeated eight times)

MOTHER: Now, listen carefully, say « Nobody likes me. »

CHILD: Oh, nobody don't *likes* me.

Le but de l'exemple est de montrer que, malgré le fait que l'enfant finisse (pas aujourd'hui, mais un jour) par dire à son tour ce que sa mère lui dit, à ce moment la grammaire mentale de l'enfant est à un stade où l'énoncé « correct » de la mère pourrait au mieux être imité, pas véritablement généré. Selon les linguistes, la mère est aussi bête que si elle s'attendait à ce que l'enfant joue une sonate pour violon ou résolve un problème de calcul, et elle reçoit ce qu'elle mérite quand ses efforts finissent par aggraver l'erreur. L'enfant joue le rôle de Laurel, et sa mère est Hardy.

De tels exemples sont puissants car ils font partie de nos propres expériences, que nous oublions jusqu'à ce qu'elles nous soient signalées, après quoi nous pouvons les essayer par nous-mêmes et obtenir généralement le même résultat. Il est extraordinaire qu'aucun des nombreux linguistes qui ont si souvent reproduit ce morceau de dialogue ne semble s'être soucié de la parentalité traumatisante qui apparaît ici : la réponse de la mère devrait être « Mon chéri, bien sûr, tu es aimé – moi, je t'aime », au lieu de renforcer la faible estime de soi du pauvre enfant.

Je ne doute pas que cette mère ait voulu le meilleur pour son enfant. Halliday et Hasan voulaient le meilleur pour Nigel, Cayla pour ses étudiants, Bernstein et Halliday pour la classe ouvrière et Chomsky pour nous, les locuteurs infiniment créatifs, avec si peu de contrôle de notre interprétation de ce qu'on nous dit que notre accord est facilement fabriqué. Ici il rejoint Ogden et Richards : leur double projet

---

14. La conversation est tirée de McNeill (1966 : 69). L'âge de l'enfant n'est pas spécifié. Joseph (2020) analyse ce passage dans un autre contexte.

du *Sens du sens* et du Basic English était motivé par leur conviction que réduire la langue à son essence éliminerait la possibilité de mentir – c'était à la suite de la Première Guerre Mondiale et de la paranoïa de la propagande qui aurait attisé les flammes de la guerre (Joseph 2002 : 181-196). Ogden et Richards imaginaient, comme Chomsky, que la production détermine l'interprétation. Ce sera George Orwell (Eric Arthur Blair, 1903-1950), avec son traitement satirique du Basic English sous la forme de Newspeak (Novlangue) dans son roman *1984*, qui fera valoir l'argument que seule la possibilité de reformuler ce qu'on nous dit en nos propres mots nous permet de déterminer si l'affirmation est vraie ou fausse, dans notre intérêt ou dans celui du gouvernement qui l'énonce (Joseph 2004). Une telle reformulation demande des alternatives linguistiques, des simplifications quand l'énoncé est complexe, des complexifications quand il est simple – exactement ce que la Novlangue rendrait impossible.

Dans la mesure où la simplicité et son contraire figurent dans toutes ces démarches, ils sont liés aux questions du vrai et du faux, de l'épistémologie et de l'enseignement, de la réalité et de l'illusion. Roman Jakobson a écrit un article publié en français sous le titre « À la recherche de l'essence du langage », en anglais « *Quest for the Essence of Language* » (Jakobson 1965). C'était en 1965, la même année où *L'Homme de La Mancha*, une version musicale de *Don Quichotte*, débutait à New York, avec sa chanson à succès « *The Quest* », « La quête » ou « L'impossible rêve ». C'est aussi l'année où *Aspects de la théorie syntaxique* a paru (Chomsky 1965). L'essence du langage : un langage distillé, réduit, purifié des inessentiels. Inévitablement simplifié à un certain niveau, mais avec des complexifications conceptuelles introduites, parfois hardiment, parfois subrepticement.

Une entreprise qu'en anglais on qualifierait de « quixotic » : chevaleresque, chimérique, utopique, et, hélas, désespérée – car chaque grande simplification est suivie du retour progressif des complexités refoulées. Voici donc ma conclusion : chaque âge a besoin d'un déblaiement, afin de préparer le terrain pour une compréhension de la langue et une pédagogie linguistique qui marcheront pour cette période particulière, dans son contexte intellectuel et éducatif plus large ; un déblaiement qu'on peut *entendre*. Il faut aussi se convaincre que l'entreprise n'est pas utopique, une conviction qui prendra des

proportions dogmatiques à mesure que les complexités rentreront progressivement sur la scène. Et enfin une nouvelle génération de linguistes pourra se moquer de la naïveté de leurs prédécesseurs et de leur réductionnisme. Notre objectif en tant qu'historiens devrait être de dire : moquez-vous-en si vous voulez, mais ne manquez pas l'occasion de comprendre pourquoi ces entreprises semblaient si justes à l'époque, surtout dans le cas de Chomsky, Bernstein et Halliday. Tesnière était hélas en avance sur son temps. Sans aucun doute, tous ces hommes étaient conscients des risques qu'ils prenaient – et des astuces qu'ils utilisaient – dans leur simplification excessive de la simplicité, pour un public qui avait tellement soif de l'essence du langage qu'il en buvait à volonté, parfois en simplifiant encore plus le message.

L'erreur fondamentale de l'analyse du discours critique et de l'analyse chomskyenne de la fabrication du consentement est d'imaginer une « tyrannie du simple » comme un exercice du pouvoir allant du haut vers le bas. Pour nous, chercheurs dans un champ aussi important que celui du langage, il importe de ne pas nous envelopper confortablement dans les complexités théoriques et le jargon, mais de trouver une « essence » et de la prononcer clairement. Et finalement faire face à la venue inévitable de notre propre chevalier de la Blanche Lune.

## Bibliographie

- Arrivé, Michel. 1969. *Les Éléments de syntaxe structurale* de Lucien Tesnière. *Langue française* 1 [La syntaxe, éd. par René Lagane et Jacqueline Pinchon]. 36-40.
- Ball, Philip. 2016. The tyranny of simple explanations: The history of science has been distorted by a longstanding conviction that correct theories about nature are always the most elegant ones. *The Atlantic*, 11 Aug. [<https://www.theatlantic.com/science/archive/2016/08/occams-razor/495332/>, consultée le 29/06/2023].
- Barsky, Robert. 2011. *Zellig Harris: From American Linguistics to Social Zionism*. Cambridge, Mass. : MIT Press.
- Benveniste, Émile. 1960. Compte rendu de Tesnière (1959). *Bulletin de la Société de linguistique de Paris* 56. 20-23.

- Bernstein, Basil. 1958. Some sociological determinants of perception: An enquiry into sub-cultural differences. *British Journal of Sociology* 9. 159-174.
- Bernstein, Basil. 1959. A public language: Some sociological implications of a linguistic form. *British Journal of Sociology* 10. 311-326.
- Bernstein, Basil. 1964. Elaborated and restricted codes: Their social origins and some consequences. *American Anthropologist* 66. 55-69.
- Blommaert, Jan & Chris Bulcaen. 2000. Critical discourse analysis. *Annual Review of Anthropology* 29. 447-466.
- Chomsky, Noam. 1957. *Syntactic Structures*. La Haye : Mouton & Co.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, Mass. : MIT Press.
- Chomsky, Noam. 1970 [1968]. Remarks on nominalization. *Readings in English Transformational Grammar*, éd. par Roderick A. Jacobs & Peter S. Rosenbaum. Boston : Ginn. 184-221.
- Chomsky, Noam. 1971. *Problems of Knowledge and Freedom: The Russell Lectures*. New York : Pantheon.
- Christie, Frances. 2007. Ongoing dialogue: Functional linguistic and Bernsteinian sociological perspectives on education. *Language, Knowledge and Pedagogy : Functional Linguistic and Sociological Perspectives*, éd. par Frances Christie & J. R. Martin. Londres : Bloomsbury, 3-13.
- Ferrero, Guillaume. 1894. L'inertie mentale et la loi du moindre effort. *Revue philosophique de la France et de l'étranger* 37. 169-182.
- Foucault, Michel. 1966. *Les mots et les choses : Une archéologie des sciences humaines*. Paris : Gallimard.
- Foucault, Michel. 1984a. *Histoire de la sexualité, II : L'usage des plaisirs*. Paris : Gallimard.
- Foucault, Michel. 1984b. *Histoire de la sexualité, III : Le souci de soi*. Paris : Gallimard.
- Fromkin, Victoria A., Robert Rodman & Nina Hyams. 2014. *An Introduction to Language* [10<sup>e</sup> éd.]. Boston : Wadsworth Cengage Learning.
- Grammont, Maurice. 1938. La néophonologie. *Le Français moderne* 6. 205-211.
- Halliday, Michael A. K. 1975. *Learning How to Mean*. Londres : Arnold.
- Halliday, Michael A. K. 2004 [1975]. Learning How to Mean. *The Language of Early Childhood*, éd. par Jonathan J. Webster (Collected Works of M. A. K. Halliday, t. 4). Londres & New York : Continuum. 28-59.
- Harris, Randy Allen. 2021. *The Linguistics Wars : Chomsky, Lakoff, and the Battle over Deep Structure* [2e éd.]. Oxford : Oxford University Press.

- Jakobson, Roman. 1965. À la recherche de l'essence du langage. *Diogène* 51. 22-38. (Version anglaise : Quest for the essence of language. *Diogenes* 13. 21-37).
- Joseph, John E. 2002. *From Whitney to Chomsky : Essays in the History of American Linguistics*. Amsterdam & New York : John Benjamins.
- Joseph, John E. 2004. Créativité linguistique, interprétation et contrôle de l'esprit selon Orwell et Chomsky. *Le discours sur la langue sous les régimes autoritaires*, éd. par Patrick Sériot & Andrée Tabouret-Keller (*Cahiers de L'ILSL* 17). Lausanne. 81-92. [Réimpr. 2006. *Texto!* 11(2) : [http://www.revue-texto.net/Inedits/Joseph\\_Creativite.html](http://www.revue-texto.net/Inedits/Joseph_Creativite.html), consultée le 28/02/2023].
- Joseph, John E. 2018. *Language, Mind and Body : A Conceptual History*. Cambridge : Cambridge University Press.
- Joseph, John E. 2020. Is/Ought : Hume's guillotine, linguistics and standards of language. *Language Prescription : Values, Ideologies and Identity*, éd. par Don Chapman & Jacob D. Rawlins. Bristol : Multilingual Matters. 15-31.
- Lees, Robert B. 1957. Compte rendu de Chomsky (1957). *Language* 33(3). 375-408.
- Lees, Robert B. 1960. *The Grammar of English Nominalizations*. *International Journal of American Linguistics*, 26(3), Part 2. Bloomington, Indiana : Indiana University Research Center in Anthropology, Folklore and Linguistics (Publication 12).
- Lombroso, Cesare. 1876. *L'uomo delinquente, in rapporto all'antropologia, alla giurisprudenza ed alla psichiatria (Cause e rimedi)*. Torino : Fratelli Bocca.
- Lombroso, Cesare. 1890. Innovation and inertia in the world of psychology. *The Monist* 1. 344-361.
- Martinet, André. 1938a. La phonologie. *Le Français moderne* 6. 131-146.
- Martinet, André. 1938b. La phonologie synchronique et diachronique. *Conférences de l'Institut de linguistique de Paris* 6. 41-58.
- Martinet, André. 1952. Function, structure and sound change. *Word* 8. 1-32.
- Martinet, André. 2005 [1955]. *Économie des changements phonétiques. Traité de phonologie diachronique*. Paris : Maisonneuve et Larose.
- Mazziotta, Nicolas. 2019. The evolution of spatial rationales in Tesnière's stemmas. *Proceedings of the Fifth International Conference on Dependency Linguistics, Paris, 2019*. <https://www.aclweb.org/anthology/W19-7709.pdf>, consultée le 28/02/2023.
- McElvenny, James. 2018. *Language and Meaning in the Age of Modernism : C. K. Ogden and His Contemporaries*. Edinburgh : Edinburgh University Press.

- McNeill, David. 1966. Developmental psycholinguistics. *The Genesis of Language*, éd. par Frank Smith & George A. Miller. Cambridge, Mass. : MIT Press. 15-84.
- Newmeyer, Frederick J. 1971. The source of derived nominals in English. *Language* 47(4). 786-796.
- Peeters, Bert. 1983. L'économie dans l'enseignement d'André Martinet. *La linguistique* 19(2). 105-116.
- Swiggers, Pierre. 1994. Aux débuts de la syntaxe structurale : Tesnière et la construction d'une syntaxe. *Linguistica* 34(1). 209-219.
- Tesnière, Lucien. 1934a. Comment construire une syntaxe. *Bulletin de la Faculté des Lettres de Strasbourg*, 12<sup>e</sup> année, 7. 219-229.
- Tesnière, Lucien. 1934b. *Petite grammaire russe*. Paris : Henri Didier.
- Tesnière, Lucien. 1959. *Éléments de syntaxe structurale*. Paris : Klincksieck.
- van Dijk, Teun A. 2008. Critical discourse analysis and nominalization: Problem or pseudo-problem? *Discourse & Society* 19(6). 821-828.
- Williams, Geoff. 2015. Halliday as an international educator. *The Bloomsbury Companion to M. A. K. Halliday*, éd. par Jonathan J. Webster. Londres : Bloomsbury Academic. 327-347.
- Zipf, George K. 1935. *The Psycho-biology of Language*. Boston : Houghton Mifflin.
- Zipf, George K. 1949. *Human Behavior and the Principle of Least Effort*. Cambridge, Mass. : Addison-Wesley.



# Mesures de complexité



# The T-unit as a measure of ‘syntactic maturity’: operationalizing linguistic complexity in early generative grammar

**Margaret Thomas**

Boston College

---

## ABSTRACT

In 1964, American language educator Kellogg W. Hunt (1912–1998) proposed that the average length of what he called the ‘T-unit’ in a text could quantify its syntactic complexity. Hunt defined a T-unit as a main clause plus any modifiers or subordinated material. His research showed that average T-units increased incrementally with age in texts written by students in grades 4, 8, 12, and adult writers. Capitalizing on a mid-century vogue for linguistics, Hunt interpreted this finding in the idiom of 1960s transformational-generative grammar as indicating that writers gradually learn to impose syntactic transformations on simple sentences, resulting in more complex syntactic structures. Teachers, educational researchers, and applied linguists continue to employ the T-unit, long outliving its initial theoretical warrant.

## KEYWORDS

Hunt (Kellogg W.), measurement of syntactic complexity, T-unit, history of applied linguistics in the United States, development of writing in children

## RÉSUMÉ

En 1964, l’enseignant de langue américain Kellogg W. Hunt (1912–1998) propose que la longueur moyenne de ce qu’il appelle le « *T-unit* » dans un texte puisse quantifier sa complexité syntaxique. Hunt définit le *T-unit*

comme une proposition principale, plus tout modificateur ou matériel subordonné. Ses recherches ont montré que la longueur moyenne des *T-units* augmentait progressivement avec l'âge dans les textes des élèves de 4<sup>e</sup>, 8<sup>e</sup>, 12<sup>e</sup> année (du système scolaire des États-Unis) et chez les rédacteurs adultes. Profitant de l'engouement pour la linguistique au milieu du siècle, Hunt a interprété cette découverte dans les termes de la grammaire transformationnelle générative des années 1960, en indiquant que les rédacteurs apprenaient progressivement à utiliser des transformations syntaxiques sur des phrases simples, aboutissant ainsi à des structures de plus en plus complexes. Les enseignants, les chercheurs en éducation et en linguistique appliquée continuent de nos jours à employer le *T-unit*, qui aura donc survécu longtemps à sa justification théorique initiale.

#### MOTS-CLÉS

Hunt (Kellogg W.), mesure de la complexité syntaxique, T-unit, histoire de la linguistique appliquée aux États-Unis, développement de l'écriture chez l'enfant

---

One warrant for the importance of inquiry into linguistic complexity versus simplicity is that the notion has both the potential to be intellectually rewarding, and also the potential to have substantial social consequences. Identification of what counts as 'complex' versus 'simple' in language may carry heavy linguistic-theoretical weight. For instance, it plays a role in discussion of diachronic change, language processing, and markedness relationships. These are all issues that may shape basic conceptualizations of the nature of language. At the same time, claims about complexity versus simplicity have shaped debates over socially-significant issues that have real impact on people's lives. For example, complexity versus simplicity figure in the design of auxiliary languages, sometimes promoted as foundational to establishing international peace. They have also figured in the devaluation of pidgin and creole languages, which goes hand-in-hand with devaluation of their speakers.

In 1964, an American scholar of English and language education named Kellogg Wesley Hunt released a report promoting an instrument for quantifying syntactic complexity in written texts, which he called the 'T-unit'. Hunt based his claims for the validity of the T-unit on an analysis of English language writing samples

produced by schoolchildren, motivating it theoretically with reference to early generative grammar. Teachers and educational researchers received Hunt's T-unit enthusiastically. To some extent so did applied linguists. Other scholars expanded the scope of Hunt's work by carrying out T-unit analyses on diverse language corpora, applying it in the evaluation of spoken as well as written output, and to the speech and writing of adult foreign-language learners. By the 1980s the notion met with criticism on several counts. Scholars proposed modifications to Hunt's definition of the T-unit, or supplemented it with alternative means of quantifying complexity. Moreover, the evolution of generative grammar eroded its purported basis in linguistic theory. Nevertheless, language research in the early decades of the 21st century still sometimes employs the T-unit.

This article reviews the lifespan to date of the T-unit as a measure of linguistic complexity versus simplicity. The value in recalling its florescence lies in the story it tells about how a scholar outside of linguistics developed a measure of linguistic complexity versus simplicity and applied it to the solution of educational problems, justifying his ideas with relatively superficial references to then-novel generative grammar. The story also illustrates how for a brief period in mid-20th century America, linguistic theory achieved a reputation as a source of practical insight into language phenomena, and made itself useful to the culture at large.

## 1. Hunt's career in language studies

Kellogg Hunt (1912–1998) is not a well-known figure in the history of linguistics. In fact, he may have had little or no formal training in the scientific analysis of language. He was born and raised in rural Parkersburg, Iowa, graduating from a small local college. Hunt continued his education with a Master's degree in English from the University of Iowa, submitting a thesis on William Wordsworth in 1935. He continued on to a doctorate in 1942, completing a dissertation on William Makepeace Thackeray. After a temporary appointment at the University of Toledo, he joined the faculty of the English Department of Florida State University in 1947, where he spent the rest of his career (Anon. 1978). Hunt's publications from the 1950s to 1960 are either literary criticism (for example, essays

on Thomas Hardy or Joseph Conrad) or reflections on teaching English at the college level. Perhaps his biggest success in the first phase of his career was a 1961 volume co-edited with FSU colleague Paul Stoakes, *Our living language*. This textbook anthologizes many short literary works that routinely show up in the U.S. high school and college English curriculum, interspersed with commentary addressed to first-year college students. Selections include Eudora Welty's 'Why I live at the P.O.', Arthur Miller's 'Death of a Salesman', and Andrew Marvell's 'To his Coy Mistress'. A second edition of *Our living language*, published in 1967, opens with an enlarged collection of essays under the heading 'The language we use'. It reprints texts by critics such as C. S. Lewis and by Archibald MacLeish on stylistics, and by linguists such as Robert A. Hall, Jr. (on language standards), William G. Moulton (on the scope and goals of linguistics, under a subsection entitled 'What language is'), and Dwight Bolinger (on linguists' prioritization of speech over writing).

Expansion of the subsection 'What language is' in the second edition of *Our living language* suggests an editorial shift towards attending more to the linguistic basis of literature. Hunt's co-editor Stoakes had taught English as a foreign language abroad, in Cuba, Colombia, and Syria (Stoakes 1945); Hunt himself spent the academic year 1958–1959 as a visiting professor at the University of Damascus. Perhaps it was under Stoakes's influence that Hunt shifted his attention away from literature and pedagogy to language issues. By 1965, the date of Hunt's first major publication on the 'T-unit', he had apparently already built up a sufficient reputation for inquiry into the development of rhetorical skills that he was presented with the National Council of English Teachers' Research Award for his work on children's writing, in the first year that the award was offered. From then onward Hunt wrote almost exclusively about the development of what he called 'syntactic maturity', centered on his notion of the T-unit. In 1972, Florida State University honored him with the Robert O. Lawton Distinguished Professorship. After Hunt's death in 1998, his colleagues established an endowed Chair in English in his name.

## 2. Development of the notion of 'T-unit'

The initial goal of Hunt's research on syntactic complexity was to define a means of quantifying structural differences between the output of younger writers and that of writers with more experience. In his classic 1964 study (published as Hunt 1965a, and extended in Hunt 1966b), to which all of his subsequent publications look back, Hunt collected 1,000-word writing samples from children in the Tallahassee Florida public schools enrolled in the 4th, 8th, and 12th grades. The subject pool comprised nine boys and nine girls at each grade level, limited to those with 'average' intelligence as determined by their scores between 90 and 110 on a standardized IQ test (Hunt 1966a: 735). Students composed the writing samples in a classroom setting but without adult assistance. The texts addressed diverse topics since they were created in the normal course of instruction. For the younger students, 1,000 words spanned more than one writing assignment (Hunt 1965a: 2-4).

Hunt's publications on the T-unit supply only a few fragmentary excerpts from his corpus, with the same ones recycled in his many articles. He does not provide a sustained demonstration of how he calculated the T-unit. However, Hunt's data likely included texts that resemble the one given in (1) below, which I include as an example of the kinds of texts Hunt's youngest writers were probably producing. (1) is an unretouched in-class writing sample composed in 2003, the independent work of an eight-year-old English monolingual 4th-grade public school student. The author of the text gave it the title 'Large numbers'; his goal was to explain the procedure for estimating a numerical quantity<sup>1</sup>.

(1) The problem I had to solve how many days are there in a million seconds. I knew I had to deal with million seconds. The first step I took was to remind myself that there were sixty seconds in a minute and there sixty minutes in a hour. So I multiplied sixty by sixty by twenty-four.

---

1. The author of (1) is one of my own then elementary school-aged sons, Peter Haskin. His composition of this text predates my interest in the history of the T-unit. I exhumed it in a search through family archives for an example of writing composed under conditions as close as possible to those under which Hunt collected his data.

Our Product was eighty-six thousand, four hundred. Next we multiplied 86,400 by 10 and it was too small so we tried mutiplying it by 100 and it was too large so gradulaly we narrowed it

In addition to texts produced by a total of 54 students in 4th, 8th, and 12th grades, Hunt extended his comparison to the output of 'superior adults', that is, professional writers, by analyzing 18 1,000-word samples taken from articles published in the periodicals *Harper's* and *The Atlantic*. Hunt first removed from the students' texts any passages he considered unintelligible or syntactically incomplete, which he called 'garbles' (1964a: 11, 1970a: 4). No adult text contained a garble. Then Hunt explored various earlier-proposed measures of structural difference across texts produced by children and adults, such as sentence length or clause length. He rejected those measures one by one. Sentence boundaries in younger students' writing are not always objectively definable, because 4th graders' inexperience with the conventions of punctuation obscures what they consider to be the beginnings and ends of sentences. This undermines the basis of any measure of syntactic complexity based on the sentence. A similar measure based on the notion of a clause has to face up to the question of whether to count conjoined predicates ('Jim and I went home and rode our bikes'; Hunt 1970a: 4) as one or as two clauses.

Eventually Hunt hit upon the idea of the T-unit as an index for quantifying the growth of syntactic complexity (or 'syntactic maturity'). A T-unit comprises a 'minimal terminable unit', that is, a single main clause plus any subordinate clauses or non-clausal material embedded under it. Informally, a T-unit is the smallest word group that could be considered a grammatical sentence, regardless of how it is punctuated. Conjoined clauses, however, counted as separate T-units. Hunt's central claim is that from 4th grade to 12th grade, and beyond into full adulthood, writers use increasingly more words per T-unit. For example, compare (2) and (3):

(2) Jim and I went home and we rode our bikes

(3) Jim and I went home and rode our bikes

Hunt analyzed (2) as comprising two T-units, each five words in length, whereas he counted (3) as a single T-unit of nine words in length. The sentence in (2) is typical of early writers, who gradually transition into creating more syntactically mature sentences such as

(3). On that basis Hunt considered the T-unit a satisfactory means of measuring structural complexity in texts.

Table 1 summarizes Hunt's core data, which he cites in almost all his publications. The leftmost column records his analysis of the data in T-units, his favored measure. The first and second columns re-analyze the same data according to other proposed measures of syntactic complexity. As a point of comparison, he used the 12th graders' output as a standard against which to assess the output of other groups. For example, the 12th graders produced an average of 14.4 words per T-unit; the 8th graders produced 11.5 words, that is, 80% as many as the 12th-graders. Adult writers produced 20.3 words per T-unit, 140% as many as the 12th-graders. Compared to the number of words per sentence or per clause, Hunt considers the T-unit to be the most sensitive and most stable index of syntactic complexity. Differences across the three student groups were highly statistically significant for average length of clause and for average length of T-unit, with a correlation coefficient of .694 for average length of T-unit (Hunt 1965b: 306).

	Average length of sentences, in words	Average length of clauses, in words	Average length of T-units, in words
Grade 4 <i>n</i> = 18	13.5 (80%)	6.6 (77%)	8.6 (60%)
Grade 8 <i>n</i> = 18	15.9 (94%)	8.1 (94%)	11.5 (80%)
Grade 12 <i>n</i> = 18	16.9 (100%)	8.6 (100%)	14.4 (100%)
Statistical signif., by grade	$p < .05$	$p < .01$	$p < .01$
X <sup>2</sup> by grade	17.03	33.10	50.35
Contingency coefficient	.489	.616	.694
Superior adults <i>n</i> = 18	24.7 (147%)	11.5 (136%)	20.3 (140%)

Table 1. Lengths of sentences, clauses, and T-units in 1,000-word texts written by 3 groups of students and adult professional writers, (as a proportion of the output of Grade 12 group). Adapted from Hunt (1965b: 306)

Hunt provides few details of how he carried out his analysis. But one can attempt to reconstruct the procedure by applying it to the text in (1), entitled 'Large numbers'. After rectifying the writer's anomalous spelling, an initial step would be to remove the first six words of line 1 as a 'garble', since a missing copula renders that constituent a syntactic orphan within the first sentence. A next step would be to divide the remaining text into minimal terminal units. In two instances, that meant separating what the writer had punctuated as conjoined sentences, breaking up the third sentence into T-units 3 and 4, and distributing the multiply-conjoined last sentence into T-units 7, 8, and 9. T-unit 4 presents a minor challenge, as it is unclear whether the missing copula should cause '*and there sixty minutes in an hour*' to be discarded as a garble, or whether the word '*there*' might be an eight-year-old's variant spelling of contracted '*there*' plus '*are*' (i.e. '*there're*'). I adopted the second analysis, and accepted it as a full T-unit. After counting the words in each T-unit, the final step is to calculate the average number of words per T-unit, as shown in Table 2. For the text in (1) the mean number of words per T-unit comes to 9.4, close to Hunt's 4th-graders' mean of 8.6. Analysis of a full 1,000-word sample of the same writer's output (going beyond the 91 words of this short passage) might result in an even closer approximation of grade-appropriate T-unit length.

T-unit	No. wds. per T-unit	
1	9	<del>The problem I had to solve</del> how many days are there in a million seconds.
2	9	I knew I had to deal with million seconds.
3	17	The first step I took was to remind myself that there were sixty seconds in a minute
4	7	and there [= there're] sixty minutes in an hour.
5	8	So I multiplied sixty by sixty by twenty-four.
6	7	Our product was eighty-six thousand, four hundred.
7	6	Next we multiplied 86,400 by 10
8	12	and it was too small so we tried multiplying it by 100
9	10	and it was too large so gradually we narrowed it.

Table 2. Analysis of a text produced by a 4th-grade student, divided into T-units with spelling regularized and 'garble' in the first line crossed out, indicating word count per T-unit

This may all seem simple-minded, and in fact—it *is*. But invention of the T-unit had a large impact in educational research. There was a burst of replications and extensions of Hunt's study, which largely supported his findings, in analyses of varied data from diverse populations. Table 3 compares seven other scholars' T-unit analyses of schoolchildren's writing published in the 12 years following Hunt (1964a). Reading the data first column by column, notice the remarkable consistency of different researchers' results by grade level, so that the averages for all tested 4th graders resemble each other, as do the averages for all tested 7th graders. Notice also that, reading the data row by row, in every case where researchers analyzed the output of writers at more than one grade level, an increase in grade consistently co-occurs with an increase in average T-unit.

Study	Total <i>n</i>	Grade						
		3	4	5	6	7	8	12
Hunt (1965b)	54		8.6				11.5	14.4
Mellon (1967)	247					9.96		
O'Donnell <i>et al.</i> (1967)	30	7.67		9.34		9.99		
Blount <i>et al.</i> (1968)	64						12.3	15.0
O'Hare (1973)	83					9.66		
Veal (1974)	27		8.71		9.48			
Combs (1975)	100					9.3		
Loban (1976)	211	7.60	8.02	8.76	9.04	8.94	10.37	13.27

Table 3. Steady growth of syntactic complexity in student writing, measured in mean T-unit by grade level, across 8 studies. Adapted from Crowhurst (1979: 92)

Data like these gave the field confidence that they now had in their hands a valid means for quantifying syntactic complexity. That confidence stimulated extensions of the T-unit to measure language complexity in new contexts, among new populations. Researchers employed T-unit scores to compare the outcomes of different techniques of teaching writing (Mellon 1967; O'Hare 1973; Combs 1975), and to determine whether subjective measures of the overall 'quality' of writing correlated with Hunt's notion of syntactic maturity (Veal 1974). Applied linguists carried out T-unit analyses on prose produced by second-language learners, both adults and children (Bardovi-Harlig 1992). Those data were useful not only to classroom teachers, but to researchers looking for a way to classify study participants into homogeneous groups by proficiency (Gaies 1980), and to school administrators as a confirmation of placement decisions (Larsen-Freeman 1978). The T-unit came to be employed in evaluating not only writing, but spontaneous speech, especially by second-language learners (Foster *et al.* 2000; Iwashita 2010). It was used to quantify structural complexity in the writing of varied additional sub-populations of schoolchildren, including those identified as speakers of specific dialects (Smith *et al.* 2001); children with a diagnosis of dyslexia (Puranik *et al.* 2007); children raised in households headed by divorced versus non-divorced adults (Crow & Ward-Lonergan 2003), and in the output of deaf children communicating through sign language (Geers *et al.* 2003). T-unit analysis was even applied in literary criticism, as in a study of the poetry of Emily Dickinson (Hallen & Shakespear 2002).

As this busy industry of applying Hunt's instrument developed, researchers and educators amplified his findings (Brett 1965), explored elaborations of the T-unit (Dixon 1970), or challenged its adequacy as a measure of linguistic complexity (Mellon 1967). Some scholars wanted to tweak the definition of the T-unit to give credit to writers who produced constructions with nested levels of embedding (DiStephano & Howie 1979; Nutter 1981; Lutkus 1987; reviewed in O'Donnell 1976). Others, especially those working with second-language populations, wanted to distinguish 'error-free' T-units, as opposed to T-units that display the morphological, lexical, or syntactic anomalies that are typical of language learners (Scott & Tucker

1974). Still others suggested that the T-unit be combined with other measures of complexity, which take into account the morphological complexion of a text (Endicott 1973; Barnwell 1988), or the density of syntactic structures (Golub & Kidder 1974; Belanger 1978). Many scholars objected to Hunt's use of a random sample of 1,000 words, on the grounds that structural complexity typically varies according to rhetorical mode: complexity is lowest in narratives, increases in descriptive texts, and rises steeply in argument or expository texts (Schmeling 1969; Crowhurst 1979). In fact, the rhetorical factor may have driven upward the score achieved by the text in (1), since the writer's struggles to count the number of days that comprise a million seconds is an exposition, not a narrative. Finally, others called attention to an issue that Hunt concedes in passing, namely that a 'syntactically mature' text according to his measure is not necessarily a clear or communicative, much less a beautiful, piece of writing (Crowhurst 1983).

Even as Hunt's notion of the T-unit sometimes met with resistance, it undeniably accomplished a lot of heavy lifting in language-educational circles. It still crops up in twenty-first-century scholarship. A Google N-gram analysis shows that frequency of the use of the term 'T-unit' increased steadily from the early 1960s to 1985 before heading downhill. However, it still retains currency in works published since 2000, such as Ortega (2003), Larsen-Freeman (2009), Beers and Nagy (2011), Vahdat *et al.* (2020), Wang and Han (2021), and many other studies cited above. A recent Forum article in the journal *Applied Linguistics* that reassessed techniques for measuring syntactic complexity referred throughout to the T-unit (Deng *et al.* 2021).

### 3. Theoretical basis of the 'T-unit'

Returning now to the claim that conceptualization of linguistic complexity versus simplicity can serve both linguistic-theoretical and practical ends, although Hunt clothed the T-unit in the theoretical garb current to the context of its invention—1960s generative grammar—his major goal was to develop a material notion of language complexity as a contribution to educational research. Hunt's stance throughout is not that of a social scientist; he does not

start with a theoretical prediction and then test it empirically. Rather, Hunt's many publications typically begin with the same seemingly tame observation, namely, that younger children produce different sentences than do older children (Hunt 1965a: 29; 1966a: 732; 1969: 5; 1970a: 1). He starts with the intuitions of a seasoned teacher, then looks for a way to measure and quantify his anecdotal observations about younger versus older writers.

Hunt's treatment of conjunction is also telling. Recall that a 'minimal terminal unit' excludes conjoined clauses, so that he divides full clauses joined with 'and' (or 'or', or 'but') into separate T-units. He does not justify this move with reference to a compelling theoretically-motivated distinction between subordination and coordination. Rather, he justifies it by observing that younger writers profusely use 'and' to join independent clauses, while older writers abandon that practice: 4th grade students conjoined three times as many clauses with 'and' as did 8th graders, and four times as many as did 12th graders (Hunt 1970a: 7). On that basis, Hunt explicitly defined the T-unit to, in effect, penalize overuse of 'and' because he found it characteristic of syntactic immaturity.

It appears, therefore, that Hunt's work began from an essentially atheoretical starting point. From there, he elaborated his conception of syntactic maturity, then as an afterthought drew on the conceptual vocabulary of generative grammar to generalize and extend his findings. In a revealing move, Hunt (1964a) declares in its extended title that the structures his research isolates will 'be analyzed by transformational methods'. In his reference to 'transformational methods' Hunt seems to mean that he will adopt language pioneered by Lees (1960) and Chomsky (1965) to conceive of the relationship between sentence pairs like (2) and (3) as due to the operation of transformational rules<sup>2</sup>. Specifically, Hunt asserts that what older

---

2. 1960s transformational generative grammar did not provide a 'method' for the specific analysis of texts, nor was it founded on close observation of the language of school-aged children. Scholars like Paula Menyuk (1961) began to pioneer the latter kind of scholarship around the time that Hunt's first publications on the T-unit appeared, starting with work on preschool-aged children. Hunt (1964a) cites Menyuk's research, but his major source for background on the language of young children seems to be the pre-generativist publications of Dorothea A. McCarthy (1930, 1954).

writers do, while 4th-grade writers do not do, is apply recursive sentence-combining transformations to kernel sentences (1964a: 56, 59–61; 1965a: 157; 1967; 1968; 1970a: 30; 1970b). That is, less mature writers conjoin clauses with 'and', while more mature writers embed clauses within each other by applying multiple transformational rules, such as 'Relative Clause Formation' or 'Nominalization'. As a result, older writers produce more words per T-unit, the hallmark of syntactic maturity in Hunt's analysis. Returning to the 4th-grade text analyzed in Table 2, a more experienced writer might have combined T-unit number 7 with T-unit number 8 to yield the more mature single T-unit 'Next we multiplied 86,400 by 10, which was too small [etc.]'. Application of that single transformation would decrease the total number of T-units by 1, and increase the average length of T-units to 10.5, seemingly vaulting this writer out of Grade 4 to approach Grade 8, according to Tables 1 and 3. Hunt suggested that introducing 4th graders to explicit sentence-combining exercises might speed up their rhetorical development by inducing an increase in T-unit (1970b: 60; 1983).

Notice, however, that Hunt's analysis of the difference between less- versus more-experienced writers 'by transformational methods' does not really bring the weight of Chomsky's novel ideas about language into play, nor does mid-twentieth century generative grammar's language of 'sentence-combining transformations' provide explicit support for Hunt's claims. In fact, generative grammar was not in the business of helping students improve their writing. Chomsky's (1965) *Aspects* model—which Hunt mentioned first in passing (1965a, 1966b, 1968) and later more extensively (1970a, 1970b, 1970c)—proposes a syntactic theory, which includes some speculations about language production, processing, and acquisition, but does not propose an ideal language pedagogy. Hunt's suggestion that schoolchildren be taught sentence-combining transformations is emblematic of his less-than-full intellectual socialization into the linguistics of the day. If more evidence were needed, in a brief article, Hunt (1970b) reveals diffuse misapprehensions: in eight short pages, he manages to misappropriate the notion of 'deep structure'; misrepresent the distinction between competence and performance; misdefine both 'acquisition' and 'development'; and mangle the form

of a syntactic tree. In a critical review, McCaig (1970) takes to task Hunt and other scholars who wrote in his wake for these errors, especially for their inaccurate understanding of how competence differs from performance and why that difference matters. Hunt (1970c) offers a defensive reply to McCaig, in which he concedes some points but still does not seem to grasp the crucial distinction. Further, he spuriously aligns ‘competence’ with ‘acquisition’ and ‘performance’ with ‘development’.

#### 4. The ‘T-unit’ in historical perspective

When Chomsky introduced transformational-generative grammar in the early 1960s, it disrupted American linguistics and stimulated neighboring disciplines, such as (in the case we have seen here) language education. There was already a precedent for the role linguistics could play in language teaching, specifically, in foreign language pedagogy. Twenty years before Hunt’s research on the T-unit began, in 1943, first the American Council of Learned Societies, and soon afterward, the U.S. Army, marshalled American linguists to develop foreign-language pedagogies that would meet the wartime need for expertise in languages spoken in both the European and Asian ‘theatres’ (Moulton 1961; Cowan 1991). Mid-century American linguists, most of whom identified themselves as ‘descriptivists’, answered the call and developed apparently successful language curricula for teaching Arabic, Burmese, Danish, Hungarian, Malay, Thai, Turkish, etc. to military personnel. Their curriculum arguably presupposed the tenets of the then-popular ‘structuralist’ conceptualization of the nature of language, and of Bloomfieldian behaviorism as a model of language learning. After the war, the Army’s Intensive Language Program proved difficult to replicate in the context of civilian higher education. Nonetheless, its reputation for efficacy during World War II raised the profile of linguistics as a discipline that could solve practical problems<sup>3</sup>.

By the early 1960s, Chomskyan linguistics began to displace structuralism. Transformational generative grammar may have

---

3. I thank Jaqueline Léon for pointing out to me the relevance of the Intensive Language Program in the backdrop to Hunt’s T-unit.

inherited a certain residual prestige from the accomplishments of the descriptivist linguistics that it replaced. To that, it added the glamour of bringing to the study of language promising novel resources from mathematics and computer science, as in Chomsky (1963)<sup>4</sup>. As evidence for a rise in the public stature of linguistics, there is the fact that when the National Council of Teachers of English established an annual research award in 1964, seven of the first ten of those awards honored work in linguistics<sup>5</sup>. As another example, prospective English teachers in the 1960s were trained on publications like *Readings in Applied English Linguistics* (Allen 1958), a kaleidoscopic anthology of short texts drawn from the writings of, among others, structuralists like Leonard Bloomfield, Benjamin Lee Whorf, Charles Fries, Archibald Hill, and, in subsequent editions, generativists like Noam Chomsky, Robert B. Lees, and Robert Stockwell. The 1964 second edition of Allen's text also incorporates a chapter by Kellogg Hunt entitled 'Improving sentence structure' (Hunt 1964b [1958]).

However, even as its public profile rose, scholars weren't always successful in specifying what difference, if any, linguistics—and in particular, generative grammar—made in their own fields. One of the first groups of scholars who investigated how to extend and apply the insights of generative grammar was psychologists. Diebold (1965) reviews the ferment that started to overtake psycholinguistics in the decade between 1954 and 1964. Impressed with the first articulations of generative grammar, and with Miller and Chomsky's

---

4. Martin-Nielsen (2010, 2011) examines the socio-political dynamics surrounding the Cold War era vogue for linguistics. In her perspective, it was not that generative grammar drove forward public interest in linguistics, but rather that generative grammar capitalized on, and was carried forward by that interest, which she attributes to political and social factors. According to Martin-Nielsen, linguistics was already poised to flourish in the years when Chomsky happened to come on the scene.

5. In 1964, Kellogg Hunt was the first winner of the David H. Russell Award for Distinguished Research in the Teaching of English, followed by William Labov (1968), Raven I. McDavid (1969), Albert H. Marckwardt (1970), Carol Chomsky (1971), Harold Allen (1973) and Roger Brown (1974). Then the tide turned: in the 46 years since 1974, only two other NCTE research awards have honored linguists: M. A. K. Halliday (1981) and Geneva Smitherman (2001) ([http://ncte.org/app/uploads/2017/06/Russell\\_Award\\_Winners.pdf](http://ncte.org/app/uploads/2017/06/Russell_Award_Winners.pdf)).

(1963: 481–483) reference in passing to research that probed (what came to be called) the ‘psychological reality of transformational grammar’, scholars such as Miller and McKean (1964), Savin and Perchonock (1965), and Fodor and Garrett (1966) tried to support (or undermine) generative theory by relating transformational complexity to sentence processing. This stream of research was eventually labeled the ‘Derivational Theory of Complexity’ (DTC). It tested whether the number of transformations entailed in the derivation of a sentence was related to variables like residual memory capacity, or speed or accuracy in the interpretation of sentence meaning. That is, would speakers less accurately assess, interpret, or remember a sentence arguably derived by ordered application of (say) six transformations, relative to a sentence derived by ordered application of three transformations? The DTC was debunked, however, when empirical studies contradicted its prediction that derivational complexity correlated with processing load. Eventually researchers acknowledged that the DTC misconstrued the specific relevance of generative grammar to sentence processing (Fodor *et al.* 1974: 319–328). Although Hunt does not cite the DTC by name, or explicitly refer to its claims about language processing, some of his publications evince familiarity with it. For example, a footnote in Hunt (1965a: 114–115) anticipates some of the problems that the DTC eventually confronted<sup>6</sup>. In the intervening years, more sophisticated work has revisited the relationship between linguistic theory and sentence processing (e.g. Hawkins 1994, 2004; *cf.* Pritchett & Whitman 1994).

The Derivational Theory of Complexity is one illustration of an early, not fully satisfying, application of transformational generative grammar. Hunt’s attempt to situate the notion of T-unit theoretically may be another such illustration. Both the DTC and Hunt’s T-unit-counting and sentence-combining projects seem to have been founded on the hope that syntactic complexity could be resolved into a straightforward, mechanical, exercise in enumeration. In both cases, generative grammar seemed at first to provide an attractive,

---

6. Fodor, Bever, and Garrett (1974) review the rise and fall of the DTC, and also review early research on child language learning from a generative perspective. They nowhere mention Hunt or the T-unit.

straightforward way to decompose something inherently complicated and multi-dimensional into units that could simply be counted. Linguistics didn't really fail educators and psycholinguists; rather, educators and psycholinguists had to abandon a mistaken notion of the scope and status of generative grammar, and reconsider their confidence that linguistics could reduce complex phenomena to an inventory of simple, quantifiable, elements. Hunt's conceptualization of linguistic complexity / simplicity in the counting of T-units had a major impact in educational circles, although it only rather tenuously participated in the linguistic theory of the day. The flourishing of the notion of the T-unit illustrates how ready Americans were in the mid 20th century to invest in what they saw as the practical applications of insights from linguistics.

## Bibliography

- Allen, Harold B. 1958. *Readings in Applied English Linguistics*. New York: Appleton-Century-Crofts [2nd ed. 1964; 3rd ed. with Michael D. Linn, 1982].
- Anon. 1978. Hunt, Kellogg Wesley. *Directory of American Scholars* (7th ed.), vol. 3, Foreign languages. 228.
- Bardovi-Harlig, Kathleen. 1992. A second look at T-unit analysis: reconsidering the sentence. *TESOL Quarterly* 26(2). 390-395.
- Barnwell, David. 1988. Some Comments on T-unit Research. *System* 16(2). 187-192.
- Beers, Scott F. & William E. Nagy. 2011. Writing development in four genres from grades three to seven: Syntactic complexity and genre differentiation. *Reading and Writing* 24(2). 183-202.
- Belanger, J. F. 1978. Calculating the syntactic density score: a mathematical problem. *Research in the Teaching of English* 12(2). 149-153.
- Blount, Nathan S., Shelby L. Johnson & Wayne C. Fredrick. 1968. *A Comparison of the Writing of Eight- and Twelfth-Grade Students* (Technical Report No. 78). Madison: University of Wisconsin Research and Development Center for Cognitive Learning [ERIC Document Reproduction Service ED 035 652].
- Brett, Sue M. 1965. A new measure of syntactic maturity. *Elementary English* 42(6). 666-668.

- Chomsky, Noam. 1963. Formal properties of grammars. *Handbook of Mathematical Psychology* (vol. 2), ed. by R. Duncan Luce, Robert R. Bush & Galanter, Eugene. New York: John Wiley. 323-418.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.
- Combs, Warren E. 1975. *Further Effects and Implications of Sentence-combining Exercises for the Secondary Language Arts Curriculum*. PhD dissertation, University of Minnesota [ERIC Document Reproduction Service ED 113 744].
- Cowan, J Milton. 1991. American linguists in peace and war. *First Person Singular II: Autobiographies by North American Scholars of the Language Sciences*, ed. by E. F. Konrad Koerner. Amsterdam & Philadelphia: John Benjamins. 67-82.
- Crow, Kristina M. & Jeannene M. Ward-Lonergan. 2003. An analysis of personal event narratives produced by school-age children. Paper presented at the annual meeting of the Council for Exceptional Children, New York, April 3-6, 2002 [ERIC Document Reproduction Service ED 481 292].
- Crowhurst, Marion. 1979. On the misinterpretation of syntactic complexity data. *English Education* 11(2). 91-97.
- Crowhurst, Marion. 1983. Syntactic complexity and writing quality: a review. *Canadian Journal of Education / Revue canadienne de l'éducation* 8(1). 1-16.
- Deng, Yaochen, Lei Lei & Dilin Liu. 2021. Calling for more consistency, refinement, and critical consideration in the use of syntactic complexity measures for writing. *Applied Linguistics* 42(5). 1021-1028.
- Diebold, A. Richard. 1965. A survey of psycholinguistic research: 1954-1964. *Psycholinguistics: A Survey of Theory and Research Problems*, ed. by Osgood, Charles E. Osgood & Thomas A. Sebeok. Bloomington: Indiana University Press. 205-275.
- DiStefano, Philip & Sherry Howie. 1979. Sentence weights: an alternative to the T-unit. *English Education* 11(2). 98-101.
- Dixon, Edward. 1970. Indexes of syntactic maturity. *Research Report of the Research Instruments Project: TRIP*, Champaign, Ill.: National Council of Teachers of English [ERIC Document Reproduction Service ED 091 748].
- Endicott, Anthony L. 1973. A proposed scale for syntactic density. *Research in the Teaching of English* 7(1). 5-12.
- Fodor, Jerry A. & Merrill F. Garrett. 1966. Some reflections on competence and performance. *Psycholinguistics Papers: The Proceedings of the 1966 Edinburgh Conference*, ed. by John Lyons & Roger J. Wales. Edinburgh: Edinburgh University Press. 135-179.

- Fodor, Jerry A. & Merrill F. Garrett. 1967. Some syntactic determinants of sentential complexity. *Perception and Psychophysics* 2. 289-296.
- Fodor, Jerry A., Thomas G. Bever & Merrill F. Garrett. 1974. *The Psychology of Language: An Introduction to Psycholinguistics and Generative Grammar*. New York: McGraw Hill.
- Foster, Pauline, Alan Tonkyn & Gillian Wigglesworth. 2000. Measuring spoken language: A unit for all reasons. *Applied Linguistics* 21(3). 354-375.
- Gaies, Stephen J. 1980. T-unit analysis in second language research: applications, problems, and limitations. *TESOL Quarterly* 14(1). 53-60.
- Geers, Ann E., Johanna G. Nicholas & Allison L. Sedey. 2003. Language skills of children with early cochlear implantation. *Ear and Hearing* 24. 46S-58S.
- Golub, Lester S. & Carole Kidder. 1974. Syntactic density and the computer. *Elementary English* 51(8). 1128-1131.
- Hallen, Cynthia L. & Jeannifer Shakespear. 2002. The T-unit as a measure of syntactic complexity in Emily Dickinson's poems. *The Emily Dickinson Journal* 111(1). 91-103.
- Hawkins, John A. 1994. *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.
- Hawkins, John A. 2004. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.
- Hunt, Kellogg W. 1964a. *Differences in Grammatical Structures Written at Three Grade Levels, the Structures to be Analyzed by Transformational Methods* (Cooperative Research Project 1998). Washington, DC: U.S. Department of Health, Education, and Welfare, Office of Education.
- Hunt, Kellogg W. 1964b [1958]. Improving sentence structure. *Readings in Applied English Linguistics* (2nd ed.), ed. by Harold B. Allen. New York: Appleton-Century-Crofts. 375-381. [Orig. in *The English Journal* 47 (April 1958). 206-211].
- Hunt, Kellogg W. 1965a. *Grammatical Structures Written at Three Grade Levels* (NCTE Research Report No. 3). Champaign, Ill.: National Council of Teachers of English [ERIC Document Reproduction Service ED 113 735].
- Hunt, Kellogg W. 1965b. A Synopsis of clause-to-sentence length factors. *The English Journal* 54(4). 300, 305-309.
- Hunt, Kellogg W. 1966a. Recent measures in syntactic development. *Elementary English* 43(7). 732-739.

- Hunt, Kellogg W. 1966b. *Sentence Structures Used by Superior Students in Grades Four and Twelve, and by Superior Adults* (Cooperative Research Project No. 5-0313). Washington, DC: U.S. Department of Health, Education, and Welfare, Office of Education [ERIC Document Reproduction Service ED 010 047].
- Hunt, Kellogg W. 1967. How little sentences grow into big ones. *New Directions in Elementary English*, ed. by Alexander Frazier. Champaign, Ill.: National Council of Teachers of English. 110-124. [ERIC Document Reproduction Service ED 033 098].
- Hunt, Kellogg W. 1968. *An Instrument to Measure Syntactic Maturity, Preliminary Version* (Grant OEG-2-6-062253-1398). Washington, DC: U.S. Department of Health, Education, and Welfare, Office of Education [ERIC Document Reproduction Service ED 020 926].
- Hunt, Kellogg W. 1969. Another probe into syntactic maturity. *Louisiana English Journal* 9(1). 5-15.
- Hunt, Kellogg W. 1970a. Syntactic maturity in schoolchildren and adults. *Monographs of the Society of Research in Child Development* 35(1). 1-67.
- Hunt, Kellogg W. 1970b. Do sentences in the second language grow like those in the first? *TESOL Quarterly* 4(3). 195-202.
- Hunt, Kellogg W. 1970c. Response to "How not to analyze the syntax of children", by Roger McCaig. *Elementary English* 47(5). 619-623.
- Hunt, Kellogg W. 1983. Sentence combining and the teaching of writing. *The Psychology of Written Language*, ed. by Margaret Martlew. New York: Wiley. 99-125.
- Hunt, Kellogg W. & James Paul Stoakes, ed. 1961. *Our Living Language*. Boston: Houghton Mifflin. (2nd ed. 1967)
- Iwashita, Noriko. 2010. Features of oral proficiency in task performance by EFL and JFL learners. *Selected Proceedings of the Second Language Research Forum*, ed. by Matthew Prior. Somerville, Mass.: Cascadilla Press. 2-47.
- Larsen-Freeman, Diane. 1978. An ESL index of development. *TESOL Quarterly* 12(2). 439-448.
- Larsen-Freeman, Diane. 2009. Adjusting expectations: the study of complexity, Accuracy, and fluency in second language acquisition. *Applied Linguistics* 30(4). 579-589.
- Lees, Robert B. 1960. *The Grammar of English Nominalizations*. Bloomington: Indiana University Research Center in Anthropology, Folklore, and Linguistics.

- Loban, Wayne. 1976. Language development: kindergarten through grade twelve (NCTE Research Report No. 18). Champaign, Ill.: National Council of Teachers of English [ERIC Document Reproduction Service ED 128 818].
- Lutkus, Alan. 1987. Problems in measuring syntactic development: T-units vs. sentence weight. *Journal of Teaching Writing* 6(1). 49-68.
- Martin-Nielsen, Janet. 2010. 'This war for men's minds': the birth of a human science in Cold War America. *History of the Human Sciences* 23(5). 131-155.
- Martin-Nielsen, Janet. 2011. A forgotten social science? Creating a place for linguistics in historical dialogue. *Journal of the History of the Behavioral Sciences* 47(2). 147-172.
- McCaig, Roger A. 1970. How not to analyze the syntax of children: a critique and a proposal. *Elementary English* 47(5). 612-618.
- McCarthy, Dorothea A. 1930. *The Language Development of the Pre-school Child* (University of Minnesota Institute of Child Welfare Monograph No. 4). Minneapolis: University of Minnesota Press.
- McCarthy, Dorothea A. 1954. Language development in children. *Manual of Child Psychology*, ed. by Leonard Carmichael. New York: Wiley. 492-630.
- Mellon, John C. 1967. Transformational sentence-combining: a method for enhancing the development of syntactic fluency in English composition (Final Report, Cooperative Research Project No. 5-8418). Cambridge, Mass.: Harvard Graduate School of Education [ERIC Document Reproduction Service ED 018 405].
- Menyuk, Paula. 1961. *Syntactic Structures in the Language of Children*. PhD dissertation, Boston University.
- Miller, George A. & Noam Chomsky. 1963. Finitary Models of Language Users. *Handbook of Mathematical Psychology* (vol. 2), ed. by R. Duncan Luce, Robert R. Bush & Eugene Galanter. New York: John Wiley. 419-491.
- Miller, George A. & Kathryn Ojemann McKean. 1964. A chronometric study of some relations between sentences. *Quarterly Journal of Experimental Psychology* 16(4). 297-308.
- Moulton, William G. 1961. Linguistics and language Teaching in the United States 1940-1960. *Trends in European and American linguistics 1930-1960*, ed. by Christine Mohrmann, Alf Sommerfelt & Joshua Whatmough. Antwerp: Spectrum. 82-109.
- Nutter, Norma. 1981. Relative merit of mean length of T-unit and sentence weight as indices of syntactic complexity in oral language. *English Education* 13(1). 17-19.

- O'Hare, Frank. 1973. Sentence-combining: improving student writing without formal grammar instruction (NCTE Research Report No. 15). Champaign, Ill.: National Council of Teachers of English [ERIC Document Reproduction Service ED 073 483].
- O'Donnell, Roy C. 1976. A critique of some indices of syntactic maturity. *Research in the Teaching of English* 10(1). 31-38.
- O'Donnell, Roy C., William J. Griffin & Raymond C. Norris. 1967. Syntax of kindergarten and elementary school children (NCTE Research Report No. 8). Champaign, Ill: National Council of Teachers of English.
- Ortega, Lourdes. 2003. Syntactic complexity measures and their relationship to L2 proficiency: a research synthesis of college-level L2 writing. *Applied Linguistics* 24(4). 492-518.
- Pritchett, Brad L. & John B. Whitman. 1994. Syntactic representation and interpretive Preference. *Japanese Sentence Processing*, ed. by Reiko Mazuka, & Noriko Nagai. Hillsdale, NJ: Lawrence Erlbaum. 65-76.
- Puranik, Cynthia S., Linda J. Lombardino & Lori J. Altmann. 2007. Writing through retellings: an exploratory study of language-impaired and dyslexic populations. *Reading and Writing* 20(3). 251-272.
- Savin, Harris B. & Ellen Perchonock. 1965. Grammatical structure and the immediate recall of English sentences. *Journal of Verbal Learning and Verbal Behavior* 4. 348-352.
- Schmeling, H. H. 1969. A study of the relative features and overall quality of college freshman writing. PhD dissertation, George Peabody College for Teachers. *Dissertation Abstracts International* 30. 4970A.
- Scott, Margaret Sue & G. Richard Tucker. 1974. Error analysis and English-language strategies of Arab students. *Language Learning* 24(1). 69-97.
- Smith, Tina T., Evan Lee & Hiram L. McDade. 2001. An investigation of T-units in African American English-speaking and standard American English-speaking fourth-grade children. *Communication Disorders Quarterly* 22(3). 148-157.
- Stoakes, James Paul. 1945. Wanted: teachers of English, foreign countries. *The Southwestern Social Science Quarterly* 26(2). 169-175.
- Vahdat, Sedigheh, Zohreh Gooniband Shooshtari & Mehdi Saba. 2020. Examining the association between T-unit and pausing length on the EFL perception of listening comprehension. *Journal of Teaching Language Skills* 39(2). 203-236.
- Veal, L. Ramon. 1974. Syntactic measures and rated quality in the writing of young children. *Studies in Language Education* (Report No. 8). Athens, Georgia: University of Georgia. [ERIC Document Reproduction Service ED 090 555].

Wang, Zehua & Feifei Han. 2021. Developing English language learners' oral production with a digital game-based mobile application. *Plos One* 16(1): e0232671.



# Linguistic complexity and theories of language in diachrony

**Vassilis Symeonidis & Nikolaos Lavidas**

National and Kapodistrian University of Athens

---

## ABSTRACT

Linguistic theories have approached change in linguistic complexity from different perspectives. Some studies have proposed, for instance, that complexity should be measured in absolute, language-internal terms, while other studies have favored a definition of complexity that is based on evidence provided by language-external parameters, such as frequency and processing difficulty. We propose a solution to this puzzle by introducing two distinct types of complexity: *structural* and *dynamic* complexity. We test our hypothesis on three semi-modal constructions of present-day English that are currently in the process of grammaticalization: *be going to*, *have got to* and *want to*. We argue that linguistic theory has generally ignored the interrelation of the two types of complexity and has referred to either the former or the latter type of complexity depending on the scope and the focus of each approach.

## KEYWORDS

language change, linguistic complexity, English diachrony, grammaticalization, dynamic complexity

## RÉSUMÉ

Les théories linguistiques ont abordé l'évolution de la complexité linguistique sous différents angles. Certaines études ont proposé, par exemple, de mesurer la complexité en termes absolus, internes à la langue, tandis que d'autres études ont favorisé une définition de la complexité basée sur des preuves fournies par des paramètres externes à la langue, tels que la fréquence et la difficulté de traitement. Nous proposons une solution à ce défi

en introduisant deux types distincts de complexité: la complexité *structurale* et la complexité *dynamique*. Nous testons notre hypothèse sur trois constructions semi-modales de l'anglais contemporain qui sont actuellement en cours de grammaticalisation: *be going to*, *have got to* et *want to*. Nous soutenons que la théorie linguistique a généralement ignoré l'interrelation des deux types de complexité et s'est référée soit au premier soit au second type de complexité en fonction de la portée et de l'objectif de chaque approche.

### MOTS-CLÉS

changement linguistique, complexité linguistique, diachronie de l'anglais, grammaticalisation, complexité dynamique

---

## 1. Introduction

In the present study, we examine how linguistic complexity can be estimated on the basis of the knowledge we obtain from several linguistic theories developed in different periods. Different linguistic theories have focused on different aspects of linguistic complexity, resulting therefore in different definitions of complexity and, consequently, in different conclusions. Some theories have favored an absolute definition of complexity, paying attention mostly to formal properties of language, while others understood complexity in relation to frequency and processing ease. Our attempt is to see how complexity changes diachronically; to do so, we examine the grammaticalization path of three well-studied semi-modal constructions of present-day English: *be going to*, *have got to* and *want to*. We propose a distinction between *structural complexity* and *processing difficulty*; the former is defined as the average form-to-function ratio of each construction, whereas the latter is taken to be a function of frequency parameters<sup>1</sup>. We measure each one separately

---

1. As mentioned below, frequency is one possible parameter among those processing factors that have been recognized as playing a role in the assessment of relative complexity (cf., among many others, Bybee & Thompson 1997; Langacker 2008). The positive influence of frequency and productivity on ease of processing has also been empirically tested and verified in numerous studies in the field of psycholinguistics; cf. Gerhand & Barry (1999) and Knobel, Finkbeiner & Caramazza (2008), among many others.

and then combine structural complexity with processing difficulty: we call this combination of structural complexity and processing difficulty *dynamic complexity*. We argue that linguistic theories may refer to one of the two types, not recognizing the other type of complexity and thus drawing different conclusions concerning the development of grammatical complexity in the diachrony of languages.

The paper is structured as follows: In Section 2, we refer to the different definitions of complexity proposed by different linguistic theories as well as some well-established assumptions regarding diachronic complexity. In Section 3, we mention what predictions grammaticalization theory makes about the development of complexity. Section 4 briefly presents the main characteristics of the grammaticalization cline related to the three case studies. In Section 5, we describe the development of structural and dynamic complexity that is exemplified in our case study. Finally, Section 6 summarizes the main conclusions of the present study.

## 2. Linguistic complexity in linguistic theory

Linguistic complexity has been extensively studied during the last two decades. Some past studies were concerned with comparing languages synchronically to evaluate differences in morphological, syntactic or phonological complexity (e.g., McWhorter 2001; Shosted 2006). In addition, various (psycholinguistic) experimental studies measured the relative complexity of different constructions of the same language (e.g., Gibson 1998; Hawkins 2004; Santi & Grodzinsky 2010). Finally, several studies have examined the development of grammatical complexity in the diachrony of languages; see, among others, Heine (2009) on the evolution of syntactic complexity, Koops & Hilpert (2009) on the parallel development of syntactic and pragmatic complexity, and van Trijp (2013) on the relationship between complexity and syncretism<sup>2</sup>.

Note that grammatical complexity has been a topic that has attracted the interest of many linguists and other scholars even in

---

2. See also the studies on this topic in the volume edited by Givón & Shibatani (2009).

earlier times. Until the first half of the 20<sup>th</sup> century, it was taken for granted that languages differ in complexity, and how complex a language was considered was almost always linked to how advanced or “primitive” the culture of the speakers that spoke it was perceived to be (for a history of perspectives on complexity, see Joseph & Newmeyer 2012). One of the earliest breakups from this tradition was expressed by Charles Hockett (1958: 180-181), according to whom:

[...] impressionistically it would seem that the total grammatical complexity of any language [...] is about the same as that of any other. This is not surprising, since all languages have about equally complex jobs to do [...]

The above idea reached a consensus among linguists in the last decades of the 20<sup>th</sup> century and soon gave rise to what is commonly known today as the “trade-off hypothesis”, in line with which all languages are of equal complexity but may differ as to the locus of it. In other words, one language may have more complex morphology but simpler/freer syntax (as is in fact the case in many strongly inflected languages, such as Latin or Greek), while others may exhibit poorer morphology, which is, however, compensated by a richer/stricter syntax (e.g., Modern English). The “trade-off” hypothesis was rejected more recently by McWhorter (2001), Shosted (2006) and Sinnemäki (2008), among others, mainly on the grounds that it is non-falsifiable. What we will propose in the present study can essentially be understood as a “trade-off mechanism”; albeit rather different in conception than what is generally understood under the term “trade-off” hypothesis. We will argue that, instead of different grammatical modules, it is different types of complexity that function as trade-offs of total complexity diachronically.

Moreover, several researchers have proposed that grammars show evidence of definite tendencies towards more or less complex forms diachronically. McWhorter (2001; 2007) states that natural language acquisition by children should mean a clear tendency to complexification. In contrast, in the case of languages that present high percentages of L2 learning, such as creoles or koines, there is a tendency to simplification instead, as a result of imperfect learning suggested by adult L2 acquisition. Similarly, Trudgill (2002; 2011) proposes that today’s languages undergo gradual simplification as a

result of the following general characteristics of modern languages: large numbers of speakers, less dense social networks, lower social stability and higher contact (hence, much L2 learning). Trudgill also claims that complexity tends to be maintained in “tightly-knit communities” where prescriptivism plays a more important role<sup>3</sup>. Although adult acquisition has been largely regarded as a complexity-decreasing factor, the relationship between complexity and language contact has given rise to opposite conclusions. For example, Kusters (2003), on the one hand, argues that language contact may result in simplification, and McWhorter (2001; 2007) similarly suggests that languages of wider communication (creoles, koines, etc.) are generally simpler. On the other hand, Nichols (1992) argues that languages spoken in multilingual environments are typically more complex than others.

Previous studies have approached linguistic complexity in two broad ways: either from an *absolute* point of view or from a *relative* point of view (Dahl 2004; Miestamo 2008). Absolute measures are mainly concerned with the estimation of the number of linguistic elements (morphemes, phonemes, etc.) as well as the interrelations between them: for example, how many categories are expressed grammatically in a language and the number of connections between those grammatical functions and their exponents (see Section 5.1). Given that denser interrelations make for a more complex system, languages that employ fused and/or polyexponential forms are expected to be more complex than languages with either analytical or agglutinative morphology. Furthermore, certain grammatical features always render a system more complex because they increase morphological load while offering limited information, such as *declensional allomorphy*, *agreement*, *grammatical gender*, etc. (McWhorter 2001).

An example of an absolute metric is provided by Piperski (2014), according to which structural complexity of a linguistic construction is defined as its average form-to-function correspondence ratio. Another example is McWhorter’s study (2007), which proposes the following three criteria for the estimation of “absolute”

---

3. On the relationship between complexity and prescriptivism, see also Lavidas (2018).



Instances of measures of Kolmogorov complexity in linguistics are mainly found in generative approaches. In generative grammar, complexity of a given structure is definable as equal to the complexity of structure-building operations (i.e., merge and move) required to produce it. Thus, a structure is as complex as the sum of syntactic operations that mediate the input and the output of the derivation. However, Kolmogorov complexity is formally incomputable and, according to some researchers, rather constitutes an unattainable ideal (Juola 2008: 92). Even though it has been applied to natural language, it fails to consider the dimension of complexity related to performance. It could be argued, for instance, that ‘longer’ rules may be more easily processed by speakers than ‘shorter’ ones, because they are more productive and cognitively entrenched.

In contrast to absolute approaches to complexity, a relative approach adopts a speaker-oriented performance-based position, according to which complexity is definable in terms of cost and processing difficulty. Kusters (2008: 9) defines complexity as “the amount of effort a generalized outsider has to make to become acquainted” with a language. His main interest is not if a language is complex but how complex it is to a specific group of users (speaker vs. hearer, first-language vs. second-language speakers). He examines how difficult different inflectional phenomena are for different groups. He argues, for example, that redundant agreement makes processing difficult for speakers and second-language learners, while it facilitates processing for hearers and first-language learners. Moreover, according to Hawkins (2004), grammatical complexity evolves along with preferences concerning cognitive processing of language. Hawkins’ Performance-Grammar Correspondence Hypothesis (2004: 122) states the following:

Grammars have conventionalized structures in proportion to their degree of preference in performance, as evidenced by frequency of use and ease of processing.

In this respect, the complexity of a grammatical construction is expected to diminish proportionally to its processing ease and frequency. This view is based on well-established and documented theories in the realm of functional/ cognitive linguistics. A similar idea is found in Bybee (1995) and Bybee & Thompson (1997), who

suggest that the higher type frequency of a linguistic item weakens its lexical status and makes it more productive, as the linguistic item is gradually emancipated from its context and acquires increasingly a grammatical function. Langacker (1987; 2008) also proposes that higher productivity (and higher type frequency) of a grammatical schema leads to its higher entrenchment in speakers' minds, which then contributes to its gradual unitization (i.e., weakening of its syntactic constituents that causes loss of compositionality)<sup>6</sup>. Unitization of a construction results in its easier activation, processing and retrieval from memory, as the construction is elicited holistically as one entity.

### 3. Linguistic complexity in grammaticalization theory

Grammaticalization is the process by which lexical elements acquire grammatical functions or elements with a less (but thus already somewhat) grammatical function acquire a more grammatical function (Kuryłowicz 1965; Hopper & Traugott 2003). A typical example of grammaticalization is the formation of the Greek Future particle *tha* out of the volitional verb *thelei* followed by a *na*-complement. The biclausal construction was reanalyzed as monoclausal, and later on, the construction was phonologically eroded until it was reduced to *tha* (cf., among many others, Pappas 2001, Roberts & Roussou 2003). The distinct phases of this process are given below in (2).

- (2) [thelei [na V]] > [thelei na [V]] > [the na [V]] > [tha [V]]

In cases of grammaticalization, first, more semantic distinctions, which were not grammatically expressed before, will be marked grammatically. Second, the grammaticalizing structure will be reanalyzed and will violate the 1-form-1-meaning principle, because more than one analyzable unit will express one and the same function (the functions of the parts will not add up to the

---

6. We agree with the reviewer that the cognitive role of frequency goes further back in linguistic theory: cf., for example, its central role in Hermann Paul's account of phonetic change, compositional transparency and analogical leveling (Auer 2015).

function of the whole). This is evident from the second phase of the grammaticalization of “*thelei na*”, where the grammaticalized *thelei* was not used as a lexical volitional verb anymore but, together with the particle *na*, marked the grammatical category “Future”. The construction was still analyzable into *thel-ei + na*, although its function was not derived compositionally but was ascribed to the whole construction as a unit. Both the grammatical marking of more semantic distinctions and the violation of the 1-form-1-meaning principle suggest complexification.

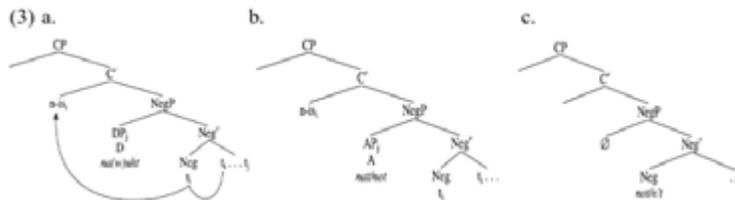
Furthermore, according to grammaticalization theory (Heine *et al.* 1991; Heine & Kuteva 2002), we also expect that the type frequency of grammaticalized items should increase due to the extension of the grammaticalized items to a wider range of contexts. As *thelei* ceased to evoke a volitional meaning in all its usages, it could appear in contexts where no such meaning was implied. Second, more frequent forms would be more easily processed and would finally become phonetically eroded. *Thelei na* was reduced to *the na*, which was then reduced to *tha*. An increase of frequency and facilitation of processing both suggest simplification.

Therefore, we face the following contradiction: grammaticalization theory predicts both *complexification* and *simplification* of grammar, depending on whether one examines complexity from an absolute or a relative point of view. Several studies have endeavored to cope with this problem by posing a division between two types of complexity. For instance, Di Sciullo (2012) and Lavidas (2018), in a diachronic study on the reorganization of grammar, have supported a distinction between I-complexity (internal), by which they mean morphosyntactic operations associated with complexity, and E-complexity (external), which is connected to frequency of use.

Grammaticalization has been thoroughly examined within a functionalist framework, but more recently it has also been investigated from a generative perspective (cf. Simpson & Wu 2002; Roberts & Roussou 2003; van Gelderen 2004a, b, 2008). Van Gelderen (2004a) claims that grammaticalization renders acquisition of a grammar easier for the language learner, because of two economy principles existing in the Minimalist program that are at work in grammaticalization. The first is the Head Preference Principle,

according to which learners prefer to build syntactic structures where functional categories are merged as heads rather than as specifiers. The Head Preference Principle can account for the preference of English speakers to use the CP head *that* to build a relative clause rather than the relative pronoun *who*, which occupies a specifier position (van Gelderen 2004a: 6). In other words, the grammaticalization of a relative pronoun to a complementizer, as well as other instances of grammaticalization (demonstratives to articles, pronouns to agreement markers, etc.), is triggered, according to van Gelderen, by this principle. The second economy principle is the Late Merge Principle, which states that it is more economical for a lexical item to merge higher in the structure at a later stage of the derivation than to merge earlier in a lower position and then move (or, in later Minimalism, internally merge) to a higher position. The Late Merge Principle accounts, for instance, for the grammaticalization of lexical verbs to auxiliaries or modals.

The two economy principles interact with the introduction of new material due to language-external reasons. The Head Preference Principle, for example, triggers the reanalysis of specifiers to heads, but new linguistic items may merge in the specifier position anew, possibly to intensify relevant features spelled out by the head. Essentially, this creates a linguistic cycle of grammaticalization for Spec-Head to Head-Head to Spec-Head again (van Gelderen 2004b: 95). Such a cycle is exemplified by the grammaticalization of negation in English demonstrated below.



(van Gelderen 2004b: 86)

In early Old English, negation was expressed by *ne* (often contracted with the verb as in 3) merging in the head of Neg and possibly moving to C, while *na(w)uht* – a reduced form of *nan wuht* (“nothing”) – could merge to Spec of NegP (3a). *Na(w)uht* reduced

further to *nat/not* in Late Middle English (3b), and by the Early Modern English period, it was reanalyzed from a specifier to a head, leaving Spec of NegP empty for new material to merge (3c). Similarly, Simpson & Wu (2002) explain the development of negation in French as a change from Move to Merge. Specifically, they suggest that, at first, negation was expressed by *ne* as the head of NegP, while the minimizer *pas* ‘step’ was base-generated inside the VP as a syntactic object and optionally moved to FocP. Later on, *pas* was reanalyzed as Spec of FocP and then as Spec of NegP, as in today’s colloquial French.

According to Roberts & Roussou (2003: 201), grammaticalization brings about simplification of structure because it causes a reordering of correspondences between features and their exponents. Roberts & Roussou (2003) propose a parametric change from Move to Merge to account for grammaticalization. They argue that, prior to grammaticalization, a single lexical item spells out features associated with two (or more) syntactic heads X and Y by externally merging in Y and then moving up the tree to X. After grammaticalization, this lexical item is reanalyzed as base-generated in X, from where it instantiates only X features. Consider again the development of the Greek Future (see 4). Before grammaticalization, *thelo*, as a lexical verb meaning ‘to want,’ takes a TP complement. *Thelo* is merged in VP and moves to TP to check phi-features. After grammaticalization, *thelo* becomes an auxiliary denoting future tense. The biclausal structure is reanalyzed as monoclausal, as *thelo* merges directly in TP and takes a VP as a complement. Therefore, the reanalysis of *thelo* from lexical verb to auxiliary is consistent with a change from Move to Merge (Roberts & Roussou 2003: 62).

- (4)  $[_{TP} \textit{to theleis} [_{VP} t_v [_{TP} \textit{mathei}]]] \rightarrow [_{TP} \textit{to theleis} [_{VP} \textit{mathei}]]$   
 (Roberts & Roussou 2003: 31)

All in all, grammaticalization in generative linguistics has been treated as a process bringing about structural simplification, which results from a preference to Merge rather than Move. Since the operation Move is better understood as Merge + Move (or double Merge), it is reasonably considered to be less economical for speakers to acquire. However, in more recent years, Move has been considered no longer as a separate structure-building operation but as a variant

of Merge (in particular, internal Merge), the most important property of which is recursion, or the possibility of the operation to apply to its own output (see, for example, Chomsky 2001). Therefore, complexity in generative linguistics has reasonably been investigated primarily in correlation to recursion. Given that recursive structures in Minimalism are seen simply as the result of iterated applications of Merge, some generative linguists have assumed that all syntactic structures must be equally complex derivationally<sup>7</sup>.

#### 4. The case studies

All three constructions examined in the present study (*be going to*, *have got to* and *want to*) constitute more or less clear cases of grammaticalization. The grammaticalization of English modals and semi-modals has been studied extensively by many researchers (cf., among many others, Krug 2000; Collins 2009; Machová 2015) who usually associate it with the emergence of phonetically eroded forms such as *gonna*, *gotta*, and *wanna* that have become increasingly frequent during the last decades (Leech *et al.* 2009: 105). In this section, we mention very briefly the grammaticalization course and the history of these constructions as well as the distinct phases we acknowledge for each one.

We start our discussion with probably the most thoroughly examined construction of the three: *be going to*. According to Traugott & König (1991), *go* in this construction began as a lexical item with an allative meaning, which later turned into a purposive marker and in more recent years into a marker of intentionality and futurity. The use of *be going to* as a Future marker spread around the late 1600s, but its first frequency boost happened during the late 1800s, and in the early 1900s the contracted formant *gonna* emerged (Danchev & Kytö 1994; Mair 2004). We distinguish three grammaticalization stages for *be going to*, or four phases in total if we take the pre-

---

7. A very recent exception to this is Bejar *et al.* (2020), who argue that linguistic complexity cannot be related exclusively to the repetitive recursion but to both processes of the computational component of the language and performance processes, an idea that had already been expressed as early as 1963 by Chomsky & Miller.

grammaticalized stage into account. These four phases are given in (5) in schematic forms. The first two phases have the same output forms, but the second is structurally reanalyzed. The third variant introduces the contracted form *gonna* accompanied by the auxiliary, whereas in the fourth the auxiliary is dropped and *gonna* is used as an “emancipated” Future marker (Lorenz 2013). It may be objected that the last variant, [*gonna* [V]], does not appear in Standard English but only in some lower prestige varieties of the language, but as long as it is documented in at least some dialects and/or sociolects of English, this observation does not affect our hypothesis.

(5) [be going [to V]] > [be going to [V]] > [be gonna [V]] > [gonna [V]]

Moving on to *have got to*, the construction exemplifies a change from possession to obligation (Bybee *et al.* 1994: 184), but its history is somewhat more obscured. It is generally accepted that the above construction is linked somehow to the development of the semantically/functionally synonymous *have to*, which adopted a meaning of modality (obligation) in Early Modern English. In the 1800s, *have got to* could be used with either a possessive or a modal meaning, presumably due to its replacing *have to* in many contexts (Mair 2012; Lorenz 2013). Four phases can be recognized here as well (see 6), with the first corresponding to the pre-grammaticalization stage. In the second phase (first phase of grammaticalization), the construction is isomorphic to the one of the first phase but is reanalyzed; the third phase consists of the auxiliary + the contracted form *gotta*, and the last phase of the contracted form, omitting the auxiliary.

(6) [*have got* [to V]] > [*have got to* [V]] > [*have gotta* [V]] > [*gotta* [V]]

Finally, *want to* demonstrates a semantic shift from necessity to volition (Krug 2000). The construction *want + to* occurred in the 1700s, denoting necessity, and later on became associated with volitional modality and gained significantly in frequency. The contracted form *wanna* is first attested in the early 1900s and is rather frequent today, albeit more restricted syntactically than *gonna* or *gotta* (Krug 2000; Lorenz 2013). It is blocked, for instance, in inflected forms of the verb (*wants to* > \**wannas*, *wanted to* > \**wannaed*) or when a silent constituent intervenes between *want* and *to*, as in (8b), where the

trace left behind by the movement of the subject (*Teddy*) blocks contraction (compare 7a, b to 8a, b).

- (7) a. *I want to succeed Teddy.*  
 b. *Teddy is the man I want to / wanna succeed <t>.*
- (8) a. *I want Teddy to succeed.*  
 b. *Teddy is the man I want <t> to / \*wanna succeed.*

(Lorenz 2013: 12)

Overall, the following three phases are recognized in the grammaticalization course of *want to*:

- (9) [*want [to V]*] > [*want to [V]*] > [*wanna [V]*]

## 5. How lessons from linguistic theories developed in different periods can apply in an examination of complexity in diachrony: structural vs. dynamic complexity

As we have mentioned in section 2, theories of complexity can roughly be divided into two main groups. On the one hand, there are theories that sustain a definition of complexity based on absolute measures, such as the number of elements in a construction and the number of interrelations between them. On the other hand, there are theories that support a relative approach to complexity in that they define complexity as a function of extra-linguistic parameters, such as frequency, ease of activation and processing. It has also been mentioned that from these two perspectives, different predictions can be made about complexity induced by language change. Some theories, for instance, consider grammaticalization to be a complexity-increasing process because it gives rise to finer semantic distinctions being grammatically marked and usually brings about more elaborate structures mediating the deep and surface levels – at least in the first stages. Other theories understand grammaticalization as a complexity-decreasing process because it facilitates mental processing for two reasons: first, by reducing the number of syntactic operations involved in the derivation, and second, by increasing the frequency and productivity of novel constructions, making them easier to activate and elicit from memory. Additionally, in order to

account for these opposing predictions, some studies distinguish between different types of complexity, such as in the cases of Di Sciullo (2012) and the diachronic study of Lavidas (2018). We acknowledge that grammaticalization predicts both complexification and simplification of grammar diachronically, and therefore we propose a similar distinction, recognizing two ways of analyzing complexity. The first is structural complexity, which constitutes an absolute measure of complexity, and the second is dynamic complexity, which combines structural complexity and frequency.

## 5. 1. Structural Complexity

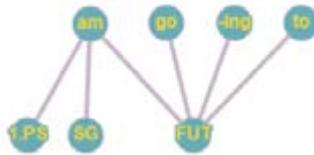
Structural complexity can be measured by using a metric system proposed by Piperski (2014) that accounts for at least two of the three criteria for absolute complexity proposed by McWhorter (2007), which are overspecification and structural elaboration. According to structural complexity, constructions can be represented as networks consisting of two tiers of vertices, one for meaning/ function and one for form, and a set of edges connecting the former to the latter. Given that each edge ( $E$ ) connects two vertices ( $n$ ), the mean number of edges incident to a vertex is equal to twice the number of the edges of a network divided by the total number of its vertices. This is also called the valency ( $v$ ) of a network<sup>8</sup>, and it is represented formally in (10). Complexity of a construction can thus be defined as the valency  $v$  of a network.

$$(10) \quad v = \frac{2E}{n}$$

For example, consider the network in Graph 1, which represents one cell of the paradigm of *be going to* at the first phase of grammaticalization. It contains 6 edges and 7 vertices in total; therefore, its valency is equal to:  $2 \times 6 / 7 = 1.71$ . In other words, on average, 1.71 edges are incident to each vertex in this network.

---

8. Piperski (2014) uses the term “average vertex degree” for the mean number of edges incident to a vertex. Here, we opted for the shorter-term “valency” to refer to the same notion.



*Graph 1:* Complexity network for one cell of the *be going to* paradigm in the first stage of grammaticalization

Hence, the total complexity of a morphological paradigm will be the sum of all networks ( $\Sigma v$ ) comprising the paradigm. Table 1 demonstrates a subset of the networks for *be going to* in every grammaticalization phase. For instance, the first network from the left represents the pre-grammaticalization stage of *be going to* (in the form of 1.PS.SG.PRES [first person, singular number, present tense]), and its average valency is relatively low ( $v=1.2$ ) because most form vertices correspond to a single function vertex with the exception of the auxiliary that encodes more information (Person, Number and Tense). In contrast, the network representing 1.PS.SG.PRES in Phase 2 (first stage of grammaticalization), which is given as an example in Graph 1, has a higher complexity ( $v=1.71$ ) because syntactic reanalysis resulted in all morphemes (form vertices) of *be going to* marking a single grammatical category (Future). More generally, networks associated with earlier stages of grammaticalization are “denser” and therefore more complex than networks linked to pre-grammaticalization stages or to later stages of grammaticalization, such as Phase 4 ( $v=1$ ), because of higher redundancy. Moreover, in Phases 1 to 3, a total of six networks can be drawn, each one corresponding to a form of the auxiliary (*am, are, is, was, were, be*), whereas only one such network exists in the last phase, since agreement features have at that point been lost following the dropping of the auxiliary<sup>9</sup>.

9. In the form-meaning networks, we did not include a vertex representing the subjects of the constructions for the following reasons. First, subjects are obligatory in all English sentences of the period we examine. Second, if we include a vertex representing the subjects, the complexity of the whole network will vary depending on how the subject position is realized (through a personal pronoun

Paradigm (indicative examples)	Phase 1	Phase 2	Phase 3	Phase 4
<i>am going to</i>				
<i>are going to</i>				

Table 1: Form-function networks of *be going to* at different phases of its grammaticalization (PS Person; SG Singular; PL Plural; PRES Present; PROG Progressive; Inf Infinitive; Fut Future)

Note that the grammaticalization of a construction rarely develops independently of the development of other constructions with similar functions. For instance, for the expression of Future tense in present-day English, there exists an array of several more or less productive variants and other constructions. Some of them are very frequent (*will*, *'ll*), others less so (*be going to*, *be gonna*) and still others much less productive and marked (*shall*). Not all of these constructions were available in all periods of English, and their frequencies were not the same as they are today. Similarly, there are several ways to express obligation, such as *must*, *have to*, *have got to*, *gotta*, etc., which developed at separate times, and their frequencies fluctuated partially due to the emergence and productivity increase/

or a noun), which is by no means determined by the semi-modal constructions. Nevertheless, we should acknowledge that the absence of a vertex for the subject impacts on the estimation of complexity at the ultimate stage of grammaticalization. The forms *gotta* and *gonna* remain essentially uninflected, and the partial overlap of features between a subject and the respective verbal form that occurs in Phases 1-3 (e.g., *she* = 3.PS + SG + FEM, *is* = 3.PS + SG + PRES) is lost in Phase 4, affecting the complexity of the respective constructions. This aspect of the problem is left to be explored in future research.

decrease of their counterparts. Although a thorough investigation of the complexity associated with the grammaticalization of a particular category (future, obligation, volition, etc.) requires the examination of all possible instantiations of the category in different periods, this goes beyond the purposes and scope of the present study. Our aim here is to explore not the complexity of whole grammatical categories but the correlations that may exist between complexity and different stages of grammaticalization of specific constructions.

Table 2 presents the structural complexity of the three constructions in all phases of grammaticalization. Numbers in black show the amount of complexity associated with each phase separately, while numbers in red show the cumulative complexity after each phase, which is how complex the system has become overall after the completion of another grammaticalization stage. This distinction is important, being based on the assumption that older variants of a construction remain in use after the emergence of novel ones, with which they are either in free variation (e.g., [*have got to* [V]] vs. [*have gotta* [V]] vs. [*gotta* [V]]) or in complementary use (e.g., [*be going to* [V]] vs. [*be going* [*to* V]]). Values of accumulated complexity after each phase of grammaticalization (numbers in red) are also plotted in Figure 1 for better visualization. As long as variation is retained, it seems that *structural complexity increases monotonously*, resulting from the creation of new grammatical features and more elaborate structures. This agrees with McWhorter (2007) and Miestamo (2008). Variants of later grammaticalization stages are of lower complexity than earlier ones, with the latest approaching a form-meaning ratio equal to 1:1 ([*gonna* [V]], [*gotta* [V]], [*wanna* V]). In contrast, form-meaning networks associated with early stages of grammaticalization are more complex, because the periphrastic constructions of these stages are often formed in non-compositional ways, which results in higher redundancy. Redundancy is reduced, however, in later stages via phonetic erosion (e.g., *go-ing to* > *gonna*).

Phase	<i>be going to</i>	<i>have got to</i>	<i>want to</i>
1 <sup>st</sup>	6.55 (6.55)	3.91 (3.91)	3.86 (3.86)
2 <sup>nd</sup>	10.22 (16.77)	4.94 (8.85)	4.70 (8.56)
3 <sup>rd</sup>	9.50 (26.27)	4.70 (13.55)	1.00 (9.56)
4 <sup>th</sup>	1.00 (27.27)	1.00 (14.55)	n/a <sup>10</sup>

Table 2: Structural complexity per phase and cumulative structural complexity

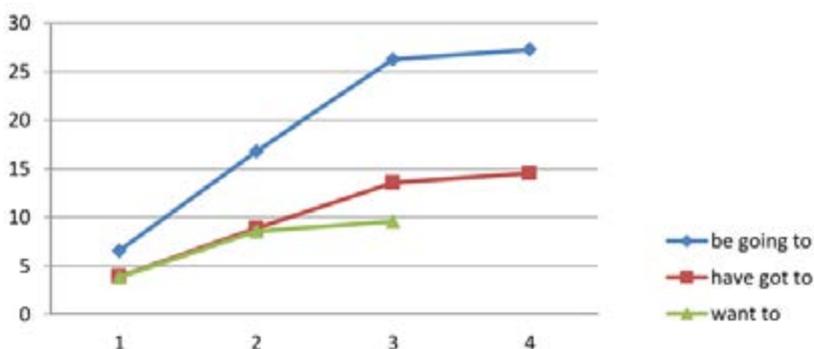


Figure 1: Cumulative Structural Complexity after each phase of grammaticalization<sup>11</sup>

## 5. 2. Dynamic Complexity

According to Hawkins (2004), a construction can be seen as simplified in proportion to its frequency. Also, for Bybee (1995: 433-435) and Langacker (2008: 17), higher frequency leads to higher entrenchment of a linguistic unit, which implies that its processing

10. Note that the 3<sup>rd</sup> phase of *want to*, [*wanna* [V]], corresponds to the 4<sup>th</sup> phase of *be going to* and *have got to*. The total number of phases for *want to* is smaller due to the fact that *want to* was not preceded by an auxiliary verb to begin with, which was lost in later phases, as was the case with the other two constructions.

11. For a detailed report of the estimation of complexity according to our methodology, see Symeonidis (2018; 2020). Here, we provide only a synopsis of our data due to limited space and because the main purposes of the present study are to emphasize the link between linguistic theory and complexity.

and activation are automated and routinized. The inverse correlation between frequency and complexity has now gained wide recognition in linguistic theory.

With regard to the case studies of the present paper, in order to gather information about type frequency of all grammaticalization phases of *be going to*, *have got to* and *want to* per decade in the period 1810-2000, we conducted simple corpus searches using the COHA corpus. We searched for the total number of occurrences of all grammaticalized variants of the three constructions followed by any verb in all decades using different search strings (8 searches in total)<sup>12</sup>. We normalized the results by calculating the relative type frequency of each construction per mil tokens in each decade. For the estimation of the type frequency of *gonna* and *gotta* (Phase 4 variants), we searched for collocations of a pronoun or noun followed by these elements in order to exclude [*V<sub>be</sub> gonna*] and [*V<sub>have</sub> gotta*], which are instances of Phase 3. Moreover, we had to remove examples that did not actually belong to Phase 4, such as questions where the inverted auxiliary appeared more than two words to the left of the contracted forms – e.g., “*Is he gonna ... ?*”.

Given that higher frequency contributes to simplification, we defined dynamic complexity (*c*) to be inversely proportionate to type frequency per mil tokens (*f*); however, at the same time, dynamic complexity increases proportionally to structural complexity as estimated by network valency (*v*). The above observation is summarized and formalized in (11) (Symeonidis 2018; 2020). The logarithm is taken here to minimize the effect of extreme values of frequency that would otherwise occur in such a wide value range. Dynamic complexity is thus a measure that considers factors of both *structure* and *frequency* simultaneously.

$$(11) \quad c = \log \frac{v}{f}$$

Figure 2 illustrates the development of dynamic complexity of the three constructions in the interval 1810-2000, using data extracted from the COHA corpus and employing the formula presented in (11).

---

12. An example of a search string we used is: *\_vh got to \_v?i*. This translates as: any form of *have* followed by *got to* followed by any verb in the infinitival form.

It is important to note here that all contracted forms (*gonna*, *gotta* and *wanna*) seem to have been attested to for the first time in about the same period (around 1910-1920) but were not very productive until one to two decades later. Therefore, we expect a significant dynamic complexity increase in this period due to the addition of three new items that were rather infrequent at the time of their emergence. From that point onward, it would seem logical that dynamic complexity decreased again following the gradual entrenchment and productivity boost of these items (cf. Hawkins 2004).

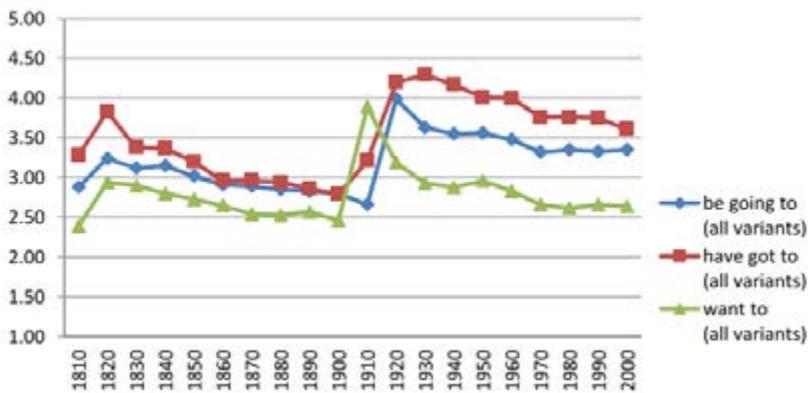


Figure 2: Dynamic Complexity per decade

Figure 2 illustrates that the monotonous upward trend found in structural complexity is canceled when frequency parameters are taken into account. Rather than a steady increase or decrease, we observe a type of *cyclical development path*: complexification by the time of emergence of new variants (around 1910-1920), followed by simplification as the new variants become increasingly more frequent. Simplification as a result of higher frequency is what Hawkins (2004) and Bybee (1995; 1997) predict. Interestingly, though, overall dynamic complexity in the interval 1980-2000 approaches the same values found in the beginning of the process, around 1810-1830. These results point towards *diachronic equicomplexity*. Moreover, the pattern of cyclicity of grammaticalization argued for by van Gelderen (2004a, b, 2008) seems to be reflected in complexity too. We should

note, though, that in the present study complexity was defined not in formal syntactic terms but as the balancing force between processing ease (as evidenced by frequency here) and structural complexity. One could possibly argue that processing ease may be the outcome of Late Merge (van Gelderen 2004) or Move to Merge (Roberts & Roussou 2003). Accordingly, the grammaticalization of *gonna*, *gotta* and *wanna* could be perceived as a change from V to T, similarly to the grammaticalization of other English modals<sup>13</sup>. However, a definition of complexity in terms of frequency has an advantage in that it treats language change as dynamic. Thus, it allows for quantitative research and renders any hypothesis concerning diachronic complexity more easily falsifiable. Nevertheless, as long as grammaticalization is understood as a process resulting in simplification due to easier processing (an assumption with which Roberts & Roussou (2003) and van Gelderen (2004) seem to agree), it is reasonable to hypothesize that any definition of processing will predict more or less the same results regarding complexity in grammaticalization.

## 6. Conclusions

In the present study, our point of departure was the observation that grammaticalization theory predicts either complexification or simplification of a linguistic item depending on the definition and point of view one assumes in order to measure complexity. If complexity is measured in absolute terms, grammaticalization is found to contribute to complexification, as it creates more elaborate structures, higher redundancy, which is provoked by syntactic reanalysis, and denser form-meaning correspondence networks. On the contrary, if complexity is measured as a function of processing ease (whether that is from a generative perspective that defines processing based on economy principles of the language faculty or

---

13. However, there is no evidence that *gonna*, *gotta* or *wanna* merge in T in present-day English, because they do not demonstrate operator properties typical of modals, such as inversion in questions, forming negation with *not/ n't* postverbally etc. (cf. Machová 2015).

from a functionalist perspective that connects it to frequency)<sup>14</sup>, we expect simplification, which can be an outcome of grammaticalization.

With the proposed metric model, we endeavored to combine these two types of complexity into what we call dynamic complexity, and we tested what happens when both structural and frequency-related factors are considered simultaneously. The results demonstrate that even though structural complexity increases as long as new characteristics and structures emerge, dynamic complexity traces a circular development path: the addition of new grammatical variants causes an increase in complexity for a short period of time, but the gradual entrenchment of the new variants and their increasing frequency cancel the increased complexity out later on reestablishing the initial complexity. More generally, it is suggested that grammars seem to become more complex diachronically when examined only in absolute terms, while they appear to approach a complexity *equilibrium* when both absolute and relative factors are considered together.

Linguistic theories differ in their approaches to complexity exactly because they examine different types of complexity. Theories examining absolute complexity predict complexification, theories investigating only phenomena related to frequency predict simplification and approaches that would analyze complexity as a function of both structure and frequency would predict equicomplexity.

## References

- Auer, Peter. 2015. Reflections on Hermann Paul as a usage-based grammarian. *Hermann Paul's Principles of Language History Revisited. Translations and Reflections*, ed by Peter Auer & Robert W. Murray. Berlin, Boston: De Gruyter. 177-207.

---

14. Note that, in this study, processing difficulty was examined only from the second perspective. That is, we defined processing difficulty in relation to frequency, not in relation to innate economy principles as it is usually understood in the generative tradition.

- Bejar, Susana, Diana Massam, Ana Teresa Pérez-Leroux & Yves Roberge. 2020. Rethinking complexity. *Syntactic Architecture and its Consequences I: Syntax inside the Grammar*, ed. by András Bány, Theresa Biberauer, Jamie Douglas, & Sten Vikner. Berlin: Language Science Press. 15-24.
- Bybee, Joan. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5). 425-455.
- Bybee, Joan & Sandra Thompson. 1997. Three frequency effects in syntax. *Annual Meeting of the Berkeley Linguistics Society* 23(1). 378-388.
- Bybee, Joan, Revere Dale Perkins & William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: The University of Chicago Press.
- Chomsky, Noam & George A. Miller. 1963. Introduction to the formal analysis of natural languages. *Handbook of Mathematical Psychology, Vol. II*, ed. by R. Duncan Luce, Robert R. Bush & Eugene Galanter. New York: John Wiley. 269-321.
- Chomsky, Noam. 2001. Derivation by phase. *Ken Hale: A Life in Language.*, ed. by Michael Kenstowicz. Cambridge, Mass.: MIT Press. 1-52.
- Collins, Peter. 2009. *Modals and Quasi-modals in English*. Amsterdam, New York: Rodopi.
- Dahl, Östen. 2004. *The Growth and Maintenance of Linguistic Complexity*. Philadelphia: John Benjamins.
- Danchev, Andrei & Merja Kytö. 1994. The construction *be going to + infinitive* in Early Modern English. *Studies in Early Modern English*, ed. by Dieter Kastovsky. Berlin: New York, Mouton de Gruyter. 59-77.
- Di Sciullo, Anna Maria. 2012. Perspectives on morphological complexity. *Current Issues in Morphology. (Ir)regularity, Frequency*, ed. by Ferenc Kiefer, Mária Ladányi & Péter Siptár. Amsterdam: John Benjamins. 105-135.
- Gerhand, Simon & Christopher Barry. 1999. Age-of-acquisition and frequency effects in speeded word naming. *Cognition* 73 (2). B27-B36.
- Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68(1). 1-76.
- Givón, Talmy & Masayoshi Shibatani, eds. 2009. *Syntactic Complexity: Diachrony, Acquisition, Neuro-cognition, Evolution*. Amsterdam: Philadelphia, John Benjamins.
- Hawkins, John A. 2004. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.
- Heine, Bernd. 2009. From nominal to clausal morphosyntax: Complexity via expansion. In Givón & Shibatani 2009. 23-51.

- Heine, Bernd, Ulrike Claudi & Friederike Hünemeyer. 1991. *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.
- Heine, Bernd & Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Hockett, Charles F. 1958. *A Course in Modern Linguistics*. New York: Macmillan.
- Hopper, Paul J. & Elizabeth Closs Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.
- Joseph, John. E. & Frederick J. Newmeyer. 2012. 'All Languages Are Equally Complex': The rise and fall of a consensus. *Historiographia Linguistica* 39(2-3). 341-368.
- Juola, Patrick. 2008. Assessing linguistic complexity. *Language Complexity: Typology, Contact, Change*, ed. by Matti Miestamo, Kaius Sinnemäki & Fred Karlsson. Amsterdam: John Benjamins. 89-108.
- Knobel, Mark, Matthew Finkbeiner & Alfonso Caramazza. 2008. The many places of frequency: Evidence for a novel locus of the lexical frequency effect in word production. *Cognitive Neuropsychology* 25(2). 256-286.
- Koops, Christian & Martin Hilpert. 2009. The co-evolution of syntactic and pragmatic complexity. Diachronic and cross-linguistic aspects of pseudoclefts. In: Givón & Shibatani 2009. 215-238.
- Krug, Manfred G. 2000. *Emerging English Modals: A Corpus-based Study of Grammaticalization*. Berlin/ New York: Mouton de Gruyter.
- Kuryłowicz, Jerzy. 1965. The evolution of grammatical categories. *Diogenes* 13. 55-71.
- Kusters, Wouter. 2003. *Linguistic Complexity. The Influence of Social Change on Verbal Inflection*. PhD Dissertation. Leiden University.
- Kusters, Wouter. 2008. Complexity in linguistic theory, language learning and language change. *Language Complexity: Typology, Contact, Change*, ed. by Matti Miestam, Kaius Sinnemäki, and Fred Karlsson. Amsterdam: John Benjamins. 3-22.
- Langacker, Ronald. W. 1987. *Foundations of Cognitive Grammar, Volume I. Theoretical Prerequisites*. Stanford: Stanford University Press.
- Langacker, Ronald W. 2008. *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press.
- Lavidas, Nikolaos. 2018. Reorganising voice in the history of Greek. *Reorganising Grammatical Variation: Diachronic Studies in the Retention, Redistribution and Refunctionalisation of Linguistic Variants*, ed. by Mirjam Schmuck, Matthias Eitelmann, & Antje Dammel. Amsterdam: John Benjamins. 175-208.

- Leech, Geoffrey, Marianne Hundt, Christian Mair & Nicholas Smith. 2009. *Change in Contemporary English: A Grammatical Study*. Cambridge: Cambridge University Press.
- Lorenz, David. 2013. *Contractions of English Semi-modals: The Emancipating Effect of Frequency*. PhD Dissertation. Albert-Ludwigs-Universität, Freiburg.
- Machová, Dagmar. 2015. *Polyfunctionality and the Ongoing History of English Modals*. PhD Dissertation. Palacký University in Olomouc.
- Mair, Christian. 2004. Corpus linguistics and grammaticalisation theory. Statistics, frequencies, and beyond. *Corpus Approaches to Grammaticalization in English*, ed. by Hans Lindquist, & Christian Mair. Amsterdam, Philadelphia: John Benjamins. 121-150.
- Mair, Christian. 2012. From opportunistic to systematic use of the web as corpus: *Do*-support with *got (to)* in contemporary American English. *The Oxford Handbook of the History of English*, ed. by Terttu Nevalainen & Elizabeth Closs Traugott. Oxford: Oxford University Press. 245-255.
- McWhorter, John. 2001. The world's simplest grammars are creole grammars. *Linguistic Typology* 5. 125-166.
- McWhorter, John. 2007. *Language Interrupted: Signs of Non-native Acquisition in Standard Language Grammars*. Oxford: Oxford University Press.
- Miestamo, Matti. 2008. Grammatical complexity in a cross-linguistic perspective. *Language Complexity: Typology, Contact, Change*, ed. by Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson. Amsterdam: John Benjamins. 23-41.
- Nichols, Johanna. 1992. *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press.
- Pappas, Panayiotis A. 2001. The microcosm of a morphological change: Variation in *thelō* + infinitive futures and *ēthela* + infinitive counterfactuals in Early Modern Greek. *Diachronica* 18(1). 59-92.
- Piperski, Alexander. 2014. An application of graph theory to linguistic complexity. *Yearbook of the Poznan Linguistic Meeting* 1(1). 89-102.
- Roberts, Ian, & Anna Roussou. 2003. *Syntactic Change: A Minimalist Approach to Grammaticalization*. Cambridge: Cambridge University Press.
- Santi, Andrea & Yosef Grodzinsky. 2010. fMRI adaptation dissociates syntactic complexity dimensions. *Neuroimage* 51(4). 1285-1293.
- Shosted, Ryan K. 2006. Correlating complexity: A typological approach. *Linguistic Typology* 10(1). 1-40.
- Simpson, Andrew & Zoe Wu. 2002. Agreement, shells, and focus. *Language*. 287-313.

- Sinnemäki, Kaius. 2008. Complexity trade-offs in core argument marking. *Language Complexity: Typology, Contact, Change*, ed. by Matti Miestam, Kaius Sinnemäki, and Fred Karlsson. Amsterdam: John Benjamins. 67-88.
- Symeonidis, Vassilis. 2018. *Grammatical Complexity in Diachrony: A Case Study in the "be going to" Construction*. Unpublished MA Thesis. Aristotle University of Thessaloniki.
- Symeonidis, Vassilis. 2020. Linguistic complexity in grammaticalization: A case study in the "be going to" construction. *The Naxos Papers. Vol. 1: On the English Diachrony*, ed. by Nikolaos Lavidas, Elly van Gelderen & Alexander Bergs. Newcastle: Cambridge Scholars Publishing. 111-121.
- Traugott, Elizabeth Closs & Ekkehard König. 1991. The semantics-pragmatics of grammaticalization revisited. *Approaches to Grammaticalization. Vol. I*, ed. by Elizabeth Closs Traugott & Bernd Heine. Amsterdam: John Benjamins. 189-218.
- Trudgill, Peter. 2002. *Sociolinguistic Variation and Change*. Washington, DC: Georgetown University Press.
- Trudgill, Peter. 2011. *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford: Oxford University Press.
- van Gelderen, Elly. 2004a. *Grammaticalization as Economy*. Amsterdam: John Benjamins.
- van Gelderen, Elly. 2004b. Economy, innovation, and prescriptivism: From Spec to Head and Head to Head. *The Journal of Comparative Germanic Linguistics* 7(1). 59-98.
- van Gelderen, Elly. 2008. Where did Late Merge go? Grammaticalization as feature economy. *Studia Linguistica* 62(3). 287-300.
- Van Trijp, Remi. 2013. Linguistic assessment criteria for explaining language change: A case study on syncretism in German definite articles. *Language Dynamics and Change* 3(1). 105-132.
- Vitanyi, Paul M.B., Frank J. Balbach, Rudi L. Cilibrasi, & Ming Li. 1997. Kolmogorov complexity and its applications, *Handbook of Theoretical Computer Science. Volume A: Algorithms and Complexity*, ed. by Jan Van Leeuwen. Amsterdam & Cambridge: Elsevier & MIT Press. 187-254.



# Table des matières

0. Introduction 11  
John E. Joseph et Chloé Laplantine

## Le naturel, le simple et le natif

1. L'évolution du concept d'ordre « naturel » à partir du xvii<sup>e</sup> siècle 29  
Gerda Hassler
2. Pourquoi l'italien serait-il la « meilleure » langue pour chanter ?  
ou : Comment la sonorité peut déterminer l'impression de  
simplicité d'une langue 55  
Claudia Schweitzer
3. L'émergence du couple terminologique « phrase simple / phrase  
complexe » dans les grammaires françaises 79  
Sophie Jollin-Bertocchi et Jacques-Philippe Saint-Gerand
4. Éloge de la naïveté dans les grammaires du français (xvi<sup>e</sup>-xvii<sup>e</sup> s.)  
Du naïf de la langue à la naïveté du style 105  
Nathalie Fournier et Cendrine Pagani-Naudet
5. Hebrew as “lingua simplicissima”: complexity in orations  
on the Hebrew language at the sixteenth-century Louvain  
Collegium Trilingue 137  
Maxime Maleux
6. Representations of linguistic simplicity in prehistoric fiction 153  
Mat Pires

## Comparer, niveler

7. Complexity or *copia*? Latin versus Greek in the sixteenth century 173  
Raf Van Rooy
8. À propos de la supériorité de l'arabe sur le persan dans un  
traité de grammaire arabe du x<sup>e</sup> siècle : apologie de la langue et  
expérience cognitive 201  
Julien Sibileau

9. Les langues d'Asie, une complexité cachée 231  
**Alice Vittrant**
10. Le défaut de « forme interne » dans les langues mandingues  
selon la psychologie des peuples de Steinthal (1867) 261  
**Jacques François**
11. La hiérarchisation des langues vue du Brésil au XIX<sup>e</sup> siècle : le  
traitement de la question par Júlio Ribeiro (1880) 291  
**Edicarlos De Aquino**

## La langue et la pensée

12. Porphyry's theory on linguistic simplicity and complexity in the  
frame of Neoplatonic teaching 315  
**Maria Chriti**
13. Generative complexity and psycholinguistics: divorce American style 335  
**Jean-Michel Fortis**

## Complexité et société

14. De Bernstein à Labov, la complexité des pratiques langagières :  
un facteur de hiérarchisation et de discrimination sociales et scolaires 377  
**Marie-Madeleine Bertucci**
15. La simplicité dans les théories syntaxiques et leurs applications  
pédagogiques dans les années 1930-1980 399  
**John E. Joseph**

## Mesures de complexité

16. The T-unit as a measure of 'syntactic maturity': operationalizing  
linguistic complexity in early generative grammar 427  
**Margaret Thomas**
17. Linguistic complexity and theories of language in diachrony 451  
**Vassilis Symeonidis et Nikolaos Lavidas**

# Société d'histoire et d'épistémologie des sciences du langage

Créée en janvier 1978, la SHESL est avant tout un lieu international de discussion et de circulation de l'information. Elle s'efforce de regrouper et de faire communiquer tous ceux qui s'intéressent à l'HISTOIRE, qui documente sur les modèles utilisés autrefois et/ou ailleurs, renseigne sur les processus d'évolution et leurs causalités ; à l'ÉPISTÉMOLOGIE, qui met en perspective critique et analyse les procédures cognitives de l'ensemble des sciences du langage (linguistique, grammaire, rhétorique, logique, pragmatique, philosophie du langage, sémiotique). Les moyens utilisés sont l'organisation de rencontres scientifiques, l'édition d'une revue (*Histoire Épistémologie Langage*) et la diffusion d'informations spécifiques à la discipline.

SHESL was founded in Paris in January 1978, to be an international centre for meetings and exchanges of ideas among researchers whose work touches on the HISTORY and EPISTEMOLOGY of the language sciences. The historical dimension involves the documentation of models from other times and traditions, and is central to an understanding of evolutionary processes and their causes. Epistemology offers critical perspectives on all aspects of the language sciences: linguistics, grammar, rhetoric, logic, pragmatics, philosophy of language, semiotics. The means adopted are the organization of scientific meetings, the publication of a journal (*Histoire, Épistémologie Langage*) and the dissemination of information specific to the field.

**COTISATION (ADHÉRENTS) 2023 : 30 €**  
**TARIF ÉTUDIANT : 20 €**

**INDIVIDUAL MEMBERS DUES 2023: 30 €**  
**STUDENT RATE: 20 €**

Moyens de paiement :

- chèque bancaire à l'ordre de la SHESL
  - en ligne sur <https://shesl.org/v2/adhesion>
  - virement bancaire au bénéfice de SHESL (préciser le nom de l'émetteur) :
- IBAN : FR76 1020 7000 5120 2110 8594 462  
BIC : CCBPFRPPMTG

Methods of payment:

- by French check made out to the order of SHESL
  - online: <https://shesl.org/v2/adhesion>
  - by bank transfer to SHESL (always mention the name of the sender):
- IBAN: FR76 1020 7000 5120 2110 8594 462  
BIC: CCBPFRPPMTG

## **SHESL**

Université Paris Cité – Case 7034  
27 rue Jean-Antoine-de-Baïf  
75 013 PARIS (France)

## ADMINISTRATION DE LA SHESL

Président : Jean-Michel Fortis  
Vice-présidente : , Anne Grondeux  
Secrétaire générale : Muriel Jorge  
Secrétaire générale adjointe : Thị Kiều Ly Phạm  
Trésorier : Pierre-Yves Testenoire  
Trésorier adjoint : Lionel Dumarty

Conseil d'administration: Émilie Aussant (Paris), Jean-Luc Chevillard (Paris), Bernard Colombat (Paris), Alejandro Diaz Villalba (Paris), Lionel Dumarty (Paris), Margherita Farina (Paris), Jean-Michel Fortis (Paris), Jacques François (Caen), Anne Grondeux (Paris), Muriel Jorge (Paris), John E. Joseph (Edinburgh), Aimée Lahaussais (Paris), James McElvenny (Siegen), Sébastien Moret (Lausanne), Thị Kiều Ly Phạm (Hanoi), Christian Puech (Paris), Pascale Rabault (Paris), Pierre-Yves Testenoire (Paris), Ekaterina Velmezova (Lausanne), Otto Zwartjes (Paris).

Comité international: Natalia Bocadorova (Moscou), Lia Formigari (Rome), Daniele Gambarara (Cosence), Eduardo Guimarães (Campinas), Gerda Haßler (Potsdam), Jukka Havu (Tampere), Douglas A. Kibbee (Urbana-Champaign), Carita Klippi (Tampere), Onno Kneepkens (Groningue), Federico Albano Leoni (Rome), Franco Lo Piparo (Palerme), Kees Meerhof (Amsterdam), Brigitte Nerlich (Nottingham), Hans Josef Niederehe (Trèves), Patrick Sériot (Lausanne), Mirko Tavoni (Pise), Talbot J. Taylor (Williamsburg), Jürgen Trabant (Berlin), Marijke van der Wal (Oegstgeest), Yuri Kleiner (Saint-Pétersbourg).

Présidente d'honneur : Lia Formigari (Rome)

## COLLECTION HEL LIVRES

Directrice : Chloé Laplantine

Comité de lecture: Sylvain Auroux (CNRS), Émilie Aussant (Université Sorbonne Nouvelle), Jean-Luc Chevillard (CNRS), Bernard Colombat (Université Paris Cité), Lionel Dumarty (CNRS), Margherita Farina (CNRS), Jean-Michel Fortis (CNRS), Jean-Marie Fournier (Université Sorbonne Nouvelle), Alessandro Garcea (Sorbonne Université), Anne Grondeux (CNRS), Jean-Patrick Guillaume (Université Sorbonne Nouvelle), Muriel Jorge (Sorbonne Université), Judith Kogel (CNRS), Chloé Laplantine (CNRS), Odile Leclercq (Sorbonne Université), Jacqueline Léon (CNRS), Christian Puech (Université Sorbonne Nouvelle), Pascale Rabault-Feuerhahn (CNRS), Sergueï Tchougounnikov (Université de Bourgogne).

# *Histoire Épistémologie Langage*

La revue *Histoire Épistémologie Langage* (HEL) paraît deux fois par an au format papier et en ligne (en libre accès sur <https://journals.openedition.org/hel/>). Les numéros antérieurs à 2020 sont disponibles sur le portail *Persée* ([www.persee.fr/collection/hel](http://www.persee.fr/collection/hel)). Les volumes imprimés peuvent être acquis auprès de la SHESL au prix de 5 € pour les numéros jusqu'en 2018, et au prix de 25 € pour les numéros plus récents.

- 1(1) 1979 Sciences du langage et métalangage
- 1(2) 1979 Ellipse et grammaire
- 2(1) 1980 Éléments d'histoire de la tradition linguistique arabe
- 2(2) 1980 Répertoire bibliographique / La coupure saussurienne / La grammaire de Montague
- 3(1) 1981 Sémantiques médiévales : cinq études sur la logique et la grammaire au Moyen Âge
- 3(2) 1981 De la grammaire à la linguistique (avec des inédits de Court de Gébelin et Rask)
- 4(1) 1982 Les idéologues et les sciences du langage
- 4(2) 1982 Statut des langues / Approches des langues à la Renaissance
- 5(1) 1983 L'ellipse grammaticale : études épistémologiques et historiques
- 5(2) 1983 La sémantique logique : problèmes d'histoire et de méthode
- 6(1) 1984 Logique et grammaire
- 6(2) 1984 Genèse du comparatisme indo-européen
- 7(1) 1985 Études sur les grammairiens grecs
- 7(2) 1985 La réflexion linguistique en Grande-Bretagne, xvii<sup>e</sup>-xviii<sup>e</sup> siècles
- 8(1) 1986 Dictionnaires, grammaires, catégories, philosophie, déchiffrement
- 8(2) 1986 Histoire des conceptions de l'énonciation
- 9(1) 1987 Les premières grammaires des vernaculaires européens
- 9(2) 1987 La tradition espagnole d'analyse linguistique
- 10(1) 1988 Stratégies théoriques
- 10(2) 1988 Antoine Meillet et la linguistique de son temps
- 11(1) 1989 Sciences du langage et recherches cognitives
- 11(2) 1989 Extension et limites des théories du langage (1880-1980)
- 12(1) 1990 Progrès et révisions
- 12(2) 1990 Grammaires médiévales
- 13(1) 1991 Épistémologie de la linguistique
- 13(2) 1991 Théories et données
- 14(1) 1992 L'adjectif : perspectives historique et typologique
- 14(2) 1992 Théories linguistiques et opérations mentales
- 15(1) 1993 Histoire de la sémantique
- 15(2) 1993 Sciences du langage et outils linguistiques
- 16(1) 1994 Actualité de Peirce
- 16(2) 1994 La grammaire des Dames
- 17(1) 1995 Théories du langage et enseignement/apprentissage des langues (fin du xix<sup>e</sup> siècle/début du xx<sup>e</sup> siècle)
- 17(2) 1995 Une familière étrangeté : la linguistique russe et soviétique
- 18(1) 1996 La linguistique de l'hébreu et des langues juives

- 18(2) 1996 L'esprit et le langage
- 19(1) 1997 Construction des théories du son (Première partie)
- 19(2) 1997 Construction des théories du son (Deuxième partie)
- 20(1) 1998 Les grammaires indiennes
- 20(2) 1998 Théories des cas
- 21(1) 1999 Linguistique des langues slaves
- 21(2) 1999 Constitution de la syntaxe
- 22(1) 2000 Horizons de la grammaire alexandrine (I)
- 22(2) 2000 Horizons de la grammaire alexandrine (II)
- 23(1) 2001 Le traitement automatique des langues
- 23(2) 2001 Dix siècles de linguistique sémitique
- 24(1) 2002 Grammaire et entités lexicales
- 24(2) 2002 Politiques linguistiques 1/2
- 25(1) 2003 Politiques linguistiques 2/2
- 25(2) 2003 Les syncatégorèmes
- 26(1) 2004 Langue et espace : retours sur l'approche cognitive
- 26(2) 2004 La linguistique baltique
- 27(1) 2005 L'autonymie
- 27(2) 2005 Autour du *De Adverbio* de Priscien
- 28(1) 2006 Histoire des idées linguistiques et horizons de rétrospection
- 28(2) 2006 Hyperlangues et fabriques de langues
- 29(1) 2007 Histoire des théories du son
- 29(2) 2007 Le naturalisme linguistique et ses désordres
- 30(1) 2008 Grammaire et mathématiques en Grèce et à Rome
- 30(2) 2008 Découverte des langues à la Renaissance
- 31(1) 2009 Mathématiques et langage
- 31(2) 2009 La nomination des langues dans l'histoire
- 32(1) 2010 Catherine II et les langues
- 32(2) 2010 SDL et psychologie à la charnière des XIX<sup>e</sup> et XX<sup>e</sup> siècles
- 33(1) 2011 Linguistique appliquée et disciplinarisation
- 33(2) 2011 Histoire des idées linguistiques et horizons de rétrospection (II)
- 34(1) 2012 La linguistique cognitive : histoire et épistémologie
- 34(2) 2012 La linguistique hispanique aujourd'hui
- 35(1) 2013 Dialectes décisifs, langues prototypiques
- 35(2) 2013 Le locuteur natif : perspectives historiques et épistémologiques
- 36(1) 2014 L'activité lexicographique dans le haut Moyen Âge latin.  
Autour du *Liber glossarum*
- 36(2) 2014 Hommage à Djamel Eddine Kouloughli
- 37(1) 2015 Le tout et ses parties. Langue, système, structure
- 37(2) 2015 Faire école en linguistique au XX<sup>e</sup> siècle : l'école de Genève
- 38(1) 2016 Une autre langue globale ? Le néerlandais comme langue  
scientifique dans l'espace extra-européen (XVII<sup>e</sup>-XIX<sup>e</sup> siècles)
- 38(2) 2016 Constitution de corpus linguistiques et pérennisation des données
- 39(1) 2017 Les langues en danger : un observatoire pertinent pour les  
théories linguistiques ?
- 39(2) 2017 La grammaire sanskrite étendue
- 40(1) 2018 Représentations et opérations dans le langage : Saussure, Bally,  
Guillaume, Benveniste, Culioli
- 40(2) 2018 La tradition linguistique arabe et l'apport des grammairiens  
arabo-andalous
- 41(1) 2019 La linguistique chinoise : influences étrangères entre XIX<sup>e</sup> et XX<sup>e</sup> siècles

- 41(2) 2019 Prescriptions en langue
- 42(1) 2020 La grammaire arabe étendue
- 42(2) 2020 Genèse, origine, récapitulation. Trần Đức Thảo face aux sciences du langage
- 43(1) 2021 La grammaire grecque étendue
- 43(2) 2021 Linguistique psychologique et sémiotique : le contexte allemand et son influence
- 44(1) 2022 Ordres et impératif : approches philosophiques et linguistiques
- 44(2) 2022 Linguistique et anthropologie au début du 20<sup>e</sup> siècle
- 45(1) 2023 Phénoménologies et théories du langage autour de Merleau-Ponty

En préparation :

- 45(2) 2023 Aux racines des dictionnaires de l'hébreu. Traduire, transcrire, transmettre

*Les Dossiers d'HEL* sont le supplément électronique de la revue. Ils sont disponibles en libre accès sur <https://shesl.org/index.php/les-dossiers-dhel>. Sont parus :

1. Wilhelm von Humboldt : éditer et lire Humboldt
2. Karl Bühler : science du langage et mémoire européenne
3. Les structuralismes linguistiques : problèmes d'historiographie comparée
5. La disciplinarisation des savoirs linguistiques : histoire et épistémologie
6. Linguistiques d'intervention : des usages socio-politiques des savoirs sur le langage et les langues
7. Lecture vernaculaire de textes classiques chinois / Reading Chinese Classical texts in the Vernacular
8. L'activité lexicographique dans le haut Moyen Âge latin
9. Écriture(s) et représentations du langage et des langues
10. Le *Liber glossarum* (s. VII-VIII) : composition, sources, réception
11. Analyse et exploitation des données de corpus linguistiques
12. Aspects historiques des grammaires portugaises et brésiliennes

Achévé d'imprimer en juillet 2023  
par Ciaco Imprimerie, Louvain-la-Neuve – [www.ciaco.com](http://www.ciaco.com)  
Imprimé en Belgique

Imprimé sur du papier ayant reçu le label écologique de l'UE  
EU Ecolabel : PT/11/002

Dépôt légal : juillet 2023

« Toutes les langues et toutes les cultures sont également complexes ! ». Cette position a été, à travers le <sup>xx</sup><sup>e</sup> siècle, la réplique des linguistes et des ethnologues aux théories jugées intenable de leurs prédécesseurs, qui avaient produit des classifications hiérarchisantes de l'humanité et des langues. Après une période d'interdit scientifique, des mesures de complexité linguistique, non suspectes de propager des idées racistes ont été de nouveau proposées.

Les contributions rassemblées dans ce volume abordent les représentations de la simplicité / complexité linguistique dans le temps long d'une histoire des idées et rendent compte d'une diversité de perspectives. On est ainsi amené à suivre les raisonnements des grammairiens et des théoriciens du langage de l'Antiquité jusqu'à la période contemporaine, en parcourant des thèmes tels l'ordre naturel, la naïveté, l'abondance, etc. Des motivations théologiques, idéologiques, pédagogiques, des critiques sociales apparaissent comme les soubassements des évaluations de simplicité ou de complexité linguistiques et des hiérarchisations. Des modèles de théories biologiques, psychologiques, philosophiques semblent encore avoir servi d'appui à la formulation de ces évaluations.

Collection HEL Livres, 3.

 Université  
Paris Cité

En couverture :

Reproduction du stemma 365 des *Éléments de syntaxe structurale* (1959) de Lucien Tesnière, figurant un passage de *Booz endormi* de Victor Hugo.

30 €  
<https://shesl.org>  
ISBN : 979-10-91587-21-1

