



**HAL**  
open science

## Le français parlé

Alain Kihm

► **To cite this version:**

Alain Kihm. Le français parlé : Esquisse de morphologie flexionnelle. Éditions Thierry Marchaisse, A paraître. hal-04502782

**HAL Id: hal-04502782**

**<https://hal.science/hal-04502782>**

Submitted on 13 Mar 2024

**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.

# Le français parlé

## Esquisse de morphologie flexionnelle

Alain Kihm (CNRS — Université Paris-Cité)

« *Freedom is the freedom to say that two plus two make four. If that is granted, all else follows.* » George Orwell

### Sommaire

Préface

1. Français parlé et français orthographique
2. Pourquoi et pour qui décrire la flexion du français parlé ?
3. Sauver l'orthographe ?
4. Comment représenter le français parlé ?
5. La morphologie flexionnelle : vue d'ensemble
6. La flexion nominale
7. La flexion adjectivale
8. La flexion pronominale
9. La flexion verbale : vue d'ensemble
10. Les formes finies du verbe
11. Les formes non-finies du verbe
12. La polarité
13. Français parlé et français orthographique : similitudes, différences, conclusion

### Préface

On estime à environ sept mille — mais on entend aussi six mille, voire cinq mille — le nombre des langues parlées sur Terre. L'imprécision tient à plusieurs causes. L'extinction en est une, la plus déplorable : chaque année, en toute discrétion le plus souvent, des langues disparaissent parce qu'elles ne sont plus transmises des parents aux enfants, ou que les jeunes, les ayant acquises, les délaissent au profit de langues qui leur offrent plus de possibilités (souvent illusoire) d'ascension sociale. Le mouvement inverse, l'éclosion de nouvelles langues à partir d'anciennes, se produit sûrement, mais il est peu documenté et peu fréquent dans le monde actuel — la dernière vague fut, semble-t-il, l'émergence des langues créoles entre le seizième et la fin du dix-neuvième siècle, dans le sillage de la première mondialisation (cf. Kihm 2005). Une autre cause qui empêche de parvenir à un compte exact est qu'il est parfois difficile de décider si l'on a affaire à deux ou plusieurs langues distinctes, ou bien à des variétés d'une même. Existe-t-il seulement des critères objectifs pour ce faire ? L'intercompréhension, qui dépend de toutes sortes de facteurs socioculturels et peut aller du total au très partiel, n'en est pas un en tout état de cause. Enfin, le dénombrement exhaustif, cartésien et la description aussi achevée que possible des langues existantes n'est pas pour bien des linguistes un objectif nécessaire de la profession. On peut s'en étonner, mais il y a à cela de bonnes raisons.

La toute première est que la linguistique, comme toute science, est un vaste domaine dont les divisions peuvent n'avoir en commun que de partager le même objet d'étude. On est linguiste si l'on consacre une partie de sa vie à décrire la structure grammaticale de langues méconnues. On l'est aussi si l'on s'attache à approfondir et affiner la description de tels aspects de la syntaxe, de la phonologie ou de la sémantique de langues par ailleurs (croit-on) bien connues comme le français ou l'anglais ; ou encore à reconstruire l'évolution et la diversification historiques des langues bantoues ; ou encore à étudier comment les enfants de deux à cinq ans acquièrent la langue à laquelle leur lieu de naissance les a exposés ; et ainsi de suite. Quant à l'objectif de décrire toutes les langues, une raison pratique s'y oppose : les linguistes descriptivistes en exercice ne sont tout simplement pas en assez grand nombre, ni les financements assez importants pour investir tous les terrains qu'il faudrait, que du reste l'état du monde rend pour beaucoup, en Afrique, en Asie, etc., de plus en plus inaccessibles, voire dangereux. Et même si le monde était ouvert et paisible, on pourrait encore invoquer une raison théorique, à savoir que toute la recherche en grammaire formelle, structurale et générative, montre que derrière leur diversité les langues humaines connues partagent un fondement structurel commun. On peut donc s'estimer justifié à considérer qu'une bonne connaissance de quelques langues suffit à les connaître toutes pour l'essentiel ou, du moins, à bâtir des hypothèses qu'on tiendra pour quasi confirmées dès lors que rien dans les données accessibles ne les infirme. Point important, qui mériterait bien des commentaires, mais qui excède les ambitions du présent ouvrage.

Maintenant, il importe que je m'explique sur le « croit-on » entre parenthèses par lequel j'ai souhaité doucher notre assurance que certaines langues seraient bien connues. Dans sa brièveté cette incise n'est en effet rien moins que la pierre angulaire du présent ouvrage, le fondement qui en justifie l'entreprise. Car qui veut explorer des langues inconnues ou méconnues n'a en réalité nul besoin de s'enfoncer dans les forêts de la Nouvelle-Guinée (où elles sont légion). On en trouve en suffisance sur le pas de notre porte. C'est seulement que, presque toujours, il faut oser soulever les voiles dont les recouvrent ces « grandes » langues dont on croit tout connaître.

Le présent ouvrage est précisément consacré à l'une de ces belles inconnues, le français parlé. Inconnue, se récriera-t-on peut-être ? La différence entre oral et écrit, n'est-elle pas disséquée et exposée dans d'innombrables études et traités, techniques ou grand public ? N'existe-t-il pas une tradition d'étude du français diversement nommé « familier », « populaire », « avancé », ou « démotique », opposé à un français qu'on dira standard ou normatif afin d'éviter au possible les jugements de valeur ? Je ne récuse pas les qualificatifs entre guillemets, en particulier le dernier qui, par son allusion à la diglossie\* grecque (*δημοτική* vs. *καθαρεύουσα* 'puriste') et par son absence de connotations possiblement douteuses, me paraît assez adéquat (cf. Massot 2008 ; Rowlett 2013). Mais considérez ceci. Premièrement, ces études, surtout lorsque destinées au plus grand public, ont souvent la prescription pour objectif plus ou moins avoué : « Dites, ne dites pas ». L'exercice est antique et du reste précieux pour ce qu'il révèle du latin parlé dont sont issues les langues romanes (cf. Väänänen 1963). (La justement fameuse « Grammaire des fautes » d'Henri Frei (1899-1980) est une exception notable à ce penchant prescriptif. On verra en quoi mon approche diverge de la sienne.) Deuxièmement, pour décider si l'on a affaire à du démotique ou à du normatif (ou « classique tardif » comme le nomme Massot), on s'attache de préférence au vocabulaire, à la prononciation (*vingt euros* avec ou sans liaison, *gageure* prononcé /gazyv/ ou /gazyv/, etc.) et, le moins anecdotique, aux constructions syntaxiques acceptables dans le premier mais rejetées par le second, ou bien imposées par le second mais inusitées dans le premier. Nulle part dans ces pages on ne lira la phrase « On en conclut quoi ? » conforme à la grammaire démotique. Le genre « traité savant » me contraint — et je m'y soumet volontiers — à n'écrire qu'en normatif, soit « Qu'en conclut-on ? ». Mais on peut en trouver de

semblables dans des dialogues entre personnages de roman. D'autre part, présentant une communication à un colloque ou invité sur quelque plateau, il est probable que j'emploierai la tournure normative, sauf émotion forte. Là est l'important : le normatif se parle, dans les situations qui le requièrent d'un commun accord, et le démotique peut s'écrire. C'est en cela qu'il n'est pas faux de parler de diglossie, au sens étymologique de « deux langues ». Et la compétence\* linguistique de la majorité des francophones inclut une connaissance plus ou moins entière des situations qui demandent l'emploi de l'une ou l'autre, ainsi qu'une conscience plus ou moins claire des pénalités qu'entraîne toute infraction à cette demande — auxquelles certains choisiront de s'exposer pour toutes sortes de raisons, tandis que d'autres ne sauront y échapper par défaut de maîtrise du normatif.

Lexique, phonologie syntaxe, il manque une composante grammaticale dans cette confrontation du démotique et du normatif : la morphologie, en particulier la morphologie flexionnelle ou flexion, c'est-à-dire la déclinaison des noms et des adjectifs en nombre (singulier / pluriel) et en genre (masculin / féminin), la déclinaison des pronoms en nombre, genre et cas (selon leur fonction grammaticale), la conjugaison des verbes en temps, mode et personne-nombre. Là, les choses sont bien différentes et la dualité démotique-normatif moins évidente. La 1<sup>ère</sup> personne du pluriel du présent de l'indicatif de *finir* /fɛnir/ est *nous finissons* /nufinisɔ̃/, celle de *partir nous partons* /nupɑ̃rtɔ̃/, point. Personne, que je sache, n'ira risquer \*/nufinɔ̃/ ou \*/nupɑ̃rtisɔ̃/. Si l'on observe des « fautes » — que j'appellerai plutôt variations — c'est pour des formes peu usitées à l'oral et d'acquisition souvent scolaire (passé simple, imparfait du subjonctif), pour des verbes « irréguliers » tels que *vêtir*, sur la conjugaison desquels on hésite,<sup>1</sup> ou pour des séries infestées d'exceptions imprévisibles, ainsi les noms et adjectifs terminés en /al/, à pluriel en /o/ ou bien invariables en nombre (à l'oral) comme la majorité des noms et adjectifs, et cela au mépris de la graphie qui embaume une flexion disparue (en tout cas inaudible) depuis le seizième siècle : cf. *(les) chats* /(le)ʃa/, *(les) festivals* /(le)festival/, etc.

Ce dernier exemple, classique, suffit à dévoiler l'originalité du français parlé en matière de flexion. Originalité pourtant méconnue, pour ne pas dire refoulée. La raison en est que la flexion de ce qu'on est convenu d'appeler le français est presque toujours décrite à partir de la forme *orthographique* des mots : le pluriel des noms et adjectifs se forme par la suffixation d'un <s> ou d'un <x> ; la 3<sup>e</sup> personne du singulier de l'imparfait de l'indicatif se termine pas un <t> qui la distingue des 1<sup>ère</sup> et 2<sup>e</sup> personnes du singulier, terminées par un <s> : *je/tu chantais* vs. *elle/il chantait*. Certes, les grammaires rédigées par des professionnels du langage n'omettent pas de signaler que ce <s> ou <x> et ce <t> ne s'entendent le plus souvent pas, ne se prononcent que dans certains contextes : liaison (*bons enfants* /bɔ̃zɑ̃fɑ̃/), inversion (*chantait-elle ?* /ʃɑ̃tɛtɛl/). Mais le dire ainsi revient à conférer à ces lettres une substance au-delà de la graphie. C'est poser que <s> ou <x> notent effectivement une composante de la flexion du français, un morphème\* de signification « pluriel » et de forme phonologique /z/ quand il se réalise — ce qu'il ne fait que si le mot suivant commence par une voyelle, et encore pas toujours.

Or, et ce sera là mon argument principal, il est bien évident que l'essentiel de la flexion est acquis et maîtrisé par les jeunes francophones avant que l'école n'entreprenne avec plus ou moins de succès de leur apprendre à lire et écrire la langue qu'elles/ils parlent déjà nativement et continueront de parler comme elles/ils l'ont acquise. Aux beaux (?) jours de la dictée scolaire, jamais, je gage, on ne les entendit imiter l'instituteur qui prononçait /lemutɔ̃s/ pour les aider à ne pas oublier d'en mettre un dans <les moutons> (cf. Cerquiligni 2007 : 68). (Il en allait ainsi a fortiori quand l'école n'était pas obligatoire en France — jusqu'en 1882 — et le nombre d'illettrés bien plus grand qu'aujourd'hui. Il en va toujours ainsi dans les pays d'Afrique dite « francophone » où le français est souvent acquis « sur le tas » par des enfants

<sup>1</sup> Kilani-Schoch & Dressler (2005 : 178-216) en donnent une liste qui a chance d'être exhaustive.

ou des adolescents peu ou pas scolarisés.) Décirer cette flexion en référence à une orthographe ignorée apparaît donc comme une entreprise scientifiquement infondée. Il convient au contraire de ne se fonder que sur ce qui s'entend, comme on y est bien forcé quand l'étude porte sur des langues non écrites.

Il me paraît donc exact d'affirmer que français démotique et français normatif *parlé* partagent pour l'essentiel, abstraction faite de phénomènes périphériques, la même flexion des noms, pronoms, adjectifs et verbes. D'où il suit que cette distinction est sans pertinence dans ce domaine. La distinction utile est entre, d'une part, le français parlé, c'est-à-dire oralisé quel qu'en soit le registre sociolinguistique ; d'autre part, le français que j'appelle « orthographique » parce que sa flexion n'a de réalité que graphique, ou bien dépourvue d'équivalent phonémique, ou bien réalisée autrement. Je désignerai désormais ces deux objets par les sigles FP et FO. La flexion du FO se lit dans les textes ; analyser la flexion du FP exige d'oublier tout à fait l'orthographe conventionnelle et de ne raisonner que sur des transcriptions phonologiques. Ce que je vais tenter de faire, bien conscient de me lancer dans une tâche pour laquelle il existe certes des précédents — notamment les travaux, anciens mais toujours actuels, de Georges Gougenheim (1900-1972), d'André Martinet (1908-1999), de Marcel Cohen (1884-1974) et ceux, fondamentaux, d'Yves Charles Morin, sans oublier le premier chapitre de Riegel et al. (1996) et l'introduction d'Abeillé & Godard (2021), en particulier Bonami (2021a) — mais somme toute peu nombreux et, surtout, le plus souvent ignorés hors du cercle des spécialistes.

Faut-il s'étonner de cette (relative) uniformité de la flexion parlée ? Non, il ne pourrait en aller autrement, me semble-t-il. Car le fond de l'affaire est sans doute que les sujets parlants ont très peu de contrôle sur la morphologie flexionnelle de leur langue, qui, avec la phonologie, en constitue véritablement le *hardware*, le matériel comme on traduit, pas trop bien. L'exemple ci-dessus de *finir* et *partir* le montre assez : on n'a pas le choix. On l'a, en revanche, entre « Qu'en conclut-on ? », « On en conclut quoi ? », « Qu'est-ce qu'on en conclut ? », voire « C'est quoi qu'on en conclut ? ». Choix contraint par divers facteurs, évidemment, mais choix quand même, car ce sont des facteurs avec lesquels on peut jouer et qui laissent une certaine marge de liberté, quand la flexion n'en laisse aucune. Même le lexique et la morphologie dérivationnelle, c'est bien connu, ne sont pas aussi contraints ; on a souvent le choix, plus ou moins conscient et là encore souvent orienté par le contexte, entre plusieurs mots synonymes, par exemple *livre* et *bouquin*, ou entre plusieurs dérivés : on dit couramment « l'invasion de l'Ukraine », mais j'ai entendu « l'envahissement de l'Ukraine », que les puristes condamneraient, bien à tort à mon avis.

En même temps, cette uniformité est en grande partie responsable de la méconnaissance de la vraie flexion du FP. Son essentielle invariance fait qu'elle ne pose en effet guère de problèmes sociolinguistiques. Qui énoncerait /nupaʁtisɔ̃/ se dénoncerait comme un étranger parlant mal le français. Il est donc facile d'en scotomiser les caractères et, vu le prestige de l'orthographe, de ne se fonder que sur elle, quitte à multiplier les avertissements : ceci s'écrit, mais ne se dit pas — bien inutiles quand adressés à des locuteurs natifs à qui on doit plutôt apprendre à ne pas oublier d'écrire un <s> là où ils n'articulent rien. Il n'est pas jusqu'à l'éloignement de la réalité phonologique qui, paradoxalement, ne favorise cette emprise de l'orthographe dans une société majoritairement lettrée. Car elle fait de l'image des mots de la langue comme un idéogramme que sa prégnance fait paraître plus réel que la parole fluctuante. *Verba volant, scripta manent* ; si l'on écrit *chats*, c'est bien qu'il y a un /s/. On verra que les linguistes n'ont pas manqué d'ingéniosité pour justifier en raison cette affirmation.

Cela étant, peut-on encore parler de diglossie ? Sans doute, mais une diglossie bien particulière, voire exceptionnelle. Comme dans la diglossie classique, les francophones doivent acquérir une compétence en deux systèmes, en l'occurrence deux systèmes

flexionnels. Mais l'un, celui du FP, n'existe qu'à l'oral, tandis que l'autre, du FO, n'existe qu'à l'écrit. Le premier s'acquiert naturellement, entre deux et cinq ans ; le second vient s'y superposer à l'âge scolaire ; ni l'un ni l'autre n'offre d'alternative : on ne peut fléchir le FP autrement, on ne peut écrire le FO autrement — ou plutôt on le pourrait en principe, mais les obstacles socioculturels sont insurmontables. Enfin, comme je vais m'attacher à le montrer, les deux systèmes sont fondamentalement différents. Situation en effet singulière où l'on parle une langue dont une partie essentielle ne s'écrit pas et où l'on en écrit une dont la même partie essentielle ne se parle pas, les deux considérées, à raison du reste, comme incarnant la même langue. Voilà une énigme digne du mystère de la Trinité. C'est ce rapprochement théologique qui m'amènera à utiliser souvent ce verbe « incarner » pour qualifier la relation du FP et du FO à la substance fictive — mais réelle et nécessaire en tant que fiction, comme l'expliquait Jeremy Bentham — qui porte le nom de « français ».

A qui ce livre s'adresse-t-il ? La France est, paraît-il, le pays où avoir chez soi un dictionnaire est la chose la plus commune. Si c'est vrai, cela indique un souci particulier de la langue. Je pourrais donc me risquer à répondre : à tout le monde ! Ce serait sans doute aller trop loin. Pourtant, le problème dont je traite est sur la place publique et régulièrement prétexte à discussions parfois acerbes. Encore récemment (début novembre 2022) les médias ont résonné d'une polémique quant à savoir ce qu'avait vraiment dit ce député du Rassemblement national qui avait crié dans l'hémicycle la suite de phonèmes /ki(l)ʁ(ə)tuʁnãnafɪk/ à propos du possible accueil d'un navire d'une ONG chargé de migrants naufragés. Était-ce, au pluriel, « Qu'ils retournent en Afrique », qui visait les migrants en personne et justifiait l'imputation de racisme ? Ou bien, comme s'en est défendu le député, au singulier, « Qu'il retourne en Afrique », « il » désignant le navire sauveteur objet du débat — mais peut-être aussi le député de gauche ainsi apostrophé, qui se trouve être d'origine africaine ? Comme on l'a fait remarquer, sur le fond le débat est sans objet : la teneur raciste du propos, de quelque façon qu'on le comprenne, n'est guère déniale. Reste que la controverse n'aurait pu avoir lieu au Parlement espagnol, par exemple : en espagnol, la différence entre les troisièmes personnes du singulier et du pluriel s'entend ; en français, elle se lit, mais ne s'entend pas toujours. Le présent ouvrage tient tout entier là. C'est pourquoi je ne désespère pas de toucher un lectorat assez large, que les questions de langue ne passionnent pas particulièrement, mais que des incidents comme celui que je viens de rapporter n'ont pu manquer d'interpeler ou d'amuser. Si je parviens à montrer à ces lectrices et lecteurs qu'il y a là bien davantage qu'une affaire d'orthographe, j'en serai satisfait.

Inutile de se le cacher, la grammaire est un exercice aride, « *the driest of all dry subjects* » estimait l'Alice de Lewis Carroll si ma mémoire est fidèle. Je ne sais si je parviendrai à la rendre attrayante. Je tâcherai en tout cas de la rendre claire. De ce point de vue, les chapitres qui composent cet ouvrage relèvent de plusieurs régimes. Les deux premiers ne requièrent, me semble-t-il, aucune connaissance particulière en matière de linguistique. La grammaire traditionnelle enseignée à l'école y suffit amplement. Les chapitres 6 à 12 constituent en revanche le cœur technique de l'étude. Je ne chercherai pas à en dissimuler l'aridité. Mais, outre que le désert est beau, ils me paraissent indispensables. Comment démontrer la profonde différence entre FO et FP sans entrer dans le détail des faits grammaticaux ? Pour ce faire, termes techniques et schématisation sont inévitables ; s'en passer tout à fait ne ferait qu'obscurcir l'exposé. Qu'on se rassure pourtant. J'en maintiendrai l'emploi au minimum nécessaire et, surtout, je n'omettrai pas de définir les premiers et d'expliquer la seconde en langage courant — ce qui est presque toujours faisable, car, si nous linguistes trouvons intérêt à utiliser un formalisme inspiré des mathématiques, notre discipline n'en partage ni l'abstraction ni la rigueur. Je n'oublie pas non plus que la morphologie flexionnelle n'est certes pas la discipline linguistique la mieux vulgarisée. L'ouvrage de Bernard Fradin (2003) est, à ma connaissance, le seul en français à en offrir une présentation

générale, mais il ne s'agit pas de vulgarisation. Aussi me suis-je risqué à la « vue d'ensemble » du chapitre 5, où je tâche d'expliquer de quoi il est question sans présupposer davantage de connaissances que celle de cette grammaire scolaire, pas si mauvaise, dont, je suppose, la grande majorité de mes lectrices et lecteurs potentiels a conservé quelque souvenir. Les outils conceptuels nécessaires à la compréhension des chapitres suivants y sont expliqués. Le glossaire en fin d'ouvrage devrait aider à rafraîchir au besoin les mémoires.

Cela étant, je vois deux publics auxquels mon travail ne peut manquer de s'adresser. Premièrement, toutes les personnes que leur activité, professionnelle ou non, confronte régulièrement à l'écart entre français écrit et français oral : professeur.e.s des écoles et des universités, journalistes, grammairien.ne.s. En second lieu, la communauté des linguistes et des étudiant.e.s en sciences du langage. Ces groupes se recoupent en partie, puisqu'il est ordinaire que professeur.e.s et grammairien.ne.s aient aussi leur casquette de linguistes. Leur relation au FP est toutefois autre. Les professeur.e.s, de lettres ou de sciences, ont pour tâche, première ou annexe, de veiller à ce que leurs élèves et étudiant.e.s fassent preuve d'une bonne maîtrise de la langue écrite, ce qui ne les contraint pas à prendre en compte la profonde différence entre FP et FO en matière de flexion. Il leur suffit de traiter cette différence comme un simple (si l'on peut dire) problème d'orthographe. Les grammairien.ne.s se consacrent à décrire le français, c'est-à-dire, presque toujours — Gougenheim (1938) une des rares exceptions — le FO. Les ouvrages de référence, tels Damourette & Pichon (1911-1927), Riegel et al. (1996), Grevisse & Goosse (2008), Grevisse (2009) ou Abeillé & Godard (2021), mentionnent certes l'oral pour l'opposer à l'écrit, mais c'est le plus souvent, délibérément ou non, sur le mode du manque, de la perte d'information. Enfin, les journalistes qui chroniquent les faits de langue n'y veulent voir en général que le pittoresque, l'amusant, le « populaire ». La prescription, la définition du « bon usage » sont communes aux premiers et aux derniers. Les grammairien.ne.s, même se réclamant de la pure description, ne parviennent guère à s'en prémunir : elles et ils disent ce qui est, les faits, on y lit ce qui doit être, la norme. Vu le rôle social de ces trois catégories de lecteurs et lectrices potentielles, il peut donc n'être pas inutile de leur rappeler (ou révéler) que le FP n'est pas un sous-produit du FO, qu'il est possible et souhaitable de le traiter, en tout cas sa flexion, comme un objet autonome, digne de l'intérêt que l'on doit à toute pratique humaine.

Je ne veux pas clore cette préface sans avoir pris une précaution, tant je suis conscient de marcher sur de la glace mince avec un tel sujet d'étude. Cette diglossie si singulière du FO et du FP, qu'on n'aille pas penser que j'en préconise ou simplement en envisage la disparition. Je la décris, je ne la dénonce pas. Puisqu'il n'est pas question de parler comme on orthographie — par exemple prononcer /lesʒats/ au lieu de /leʒa/, évidemment impensable — la réduire ne pourrait se faire qu'en orthographiant autant que possible comme on parle. Or, cela va bien au-delà d'une « réforme » ou d'une « simplification » de l'orthographe. Les diverses propositions modernes en ce sens (guère appliquées) n'ont jamais envisagé autre chose que l'orthographe dite « d'usage ».<sup>2</sup> Ainsi, selon la dernière en date, élaborée en 1990 par l'Académie française et (peut-être) mise en œuvre trente-deux ans plus tard (à la rentrée 2022), on n'écrira plus <ambiguë>, où le tréma sur le <e> le signalait comme étant la désinence du féminin, encore audible au dix-septième siècle, qu'il convenait donc de séparer du <u> du masculin notant /y/, et non partie du digramme <gu> notant /g/ devant <e>, <i> et <y>, d'où <ambigu> pour /ãbigy/ et <ambiguë> pour /ãbigyø/. (Par le fait, « tréma » se dit *diaeresis* en anglais et *diéresis* en espagnol, du grec *διαίρεσις* voulant dire « séparation ».) Mais on écrira <ambigüe> où le tréma indique désormais que le <u> se fait entendre, le <e>

---

<sup>2</sup> Je précise « modernes », car le désir de réforme orthographique ne date évidemment pas d'hier. Cerquiglini (2007 : 42) cite Louis Meigret, auteur de la première grammaire française en français, qui, en 1550, écrivait phonétiquement : « ... nou'deuons dire, come nou'dizons ». A son époque, les <s> finaux se prononçaient sous certaines conditions (cf. chapitre 6).

étant devenu « muet » en tout état de cause. Il se peut que cette nouvelle solution soit plus cohérente, vu précisément la mutité de <e>. Mais l'important, au fond, c'est le signal, où qu'il se place, puisque le but est d'éviter qu'on ne prononce /ãbig/ ce qui s'écrirait <ambigue>. Pour le reste, je vois mal en quoi une solution est plus « simple » que l'autre.

Ecrire comme on parle voudrait dire s'attaquer à l'orthographe grammaticale, par exemple supprimer les <s> de pluriel, d'où, entre autres solutions envisageables, <lé cha> au lieu de <les chats>, en omettant au passage le <t> muet. Personne (je suppose) n'y songe raisonnablement. L'image du français, plus qu'altérée, en sortirait métamorphosée, on aurait l'impression de lire une autre langue. Et on aurait raison. Car ce qu'aurait accompli le ou la révolutionnaire orthographique, c'est remplacer la flexion du FO par celle du FP, dont j'ai dit et je montrerai qu'elle est toute autre. Seul, je pense, un régime totalitaire épris de novlangue oserait s'y risquer. Il n'est pas dit qu'il y réussisse.<sup>3</sup> Il est d'ailleurs remarquable que l'Académie ait renoncé à réformer les règles d'accord du participe passé, dont tout le monde s'accorde à déplorer la complexité et le caractère artificiel, mais trop imbriquées aux faits de flexion.

Mais, s'il n'est pas question d'écrire le FP avec sa flexion — ce qui, je le répète, est bien autre chose qu'écrire du français démotique en orthographe officielle ou peu s'en faut — il me paraît néanmoins intéressant et utile de le *transcrire*, afin d'en étudier ce qui en constitue la principale spécificité, à savoir précisément cette flexion. Intérêt scientifique, je crois, mais utilité sociale aussi. Combien de francophones, quel que soit leur niveau d'instruction, n'oublent jamais de « mettre un s » à un nom ou adjectif évidemment pluriels ou d'accorder un verbe à la 3<sup>e</sup> personne du pluriel (« les oiseaux chante » au lieu de « les oiseaux chantent ») ? Je m'y suis surpris quelques lignes plus haut ! Et mon ordinateur n'a pas plus réagi alors qu'à la ligne ci-dessus, car l'orthographe grammaticale est, comme on sait, un domaine où les correcteurs automatiques se révèlent encore peu performants. J'ai donc commis une « faute ». Le mot est lourd, que le Robert définit comme « manquement à la règle morale, au devoir, mauvaise action ». Mais ma faute et l'opprobre, modéré mais cuisant, qui s'y attache éventuellement ne sont tels que parce que l'écart flexionnel entre FP et FO est méconnu — c'est-à-dire connu mais forclos — et que l'écrit conforme au premier est ressenti comme une mutilation du second. Comme si le FO représentait le vrai français, et le FP sa réalisation malapprise.

La réalité, je l'ai dit, est à l'inverse. Les jeunes francophones acquérant naturellement leur langue assimilent la flexion du FP et apprennent, par exemple, à dire /lãfamjol/ ou /leʃamjol/ selon qu'un ou plusieurs exemplaires de *Felis catus* réclame(nt) sa/leur pitance. C'est de cette flexion — nombre marqué par l'alternance du déterminant, verbe invariable pour certaines valeurs — qu'elles et ils conserveront la compétence active toute leur vie. Puis, à un certain moment de leur existence, on leur inculque un autre système flexionnel, graphique, en vertu duquel on doit écrire <le chat miaule> ou <leṣ chatṣ miaulent> pour décrire les mêmes situations. L'expérience démontre, me semble-t-il, que l'assimilation de cet autre système, profondément différent, peut être satisfaisante, voire excellente, mais jamais parfaite. Ce ne sera jamais qu'un greffon plus ou moins réussi sur le tronc de la compétence native.

Dans ces conditions, les erreurs d'orthographe grammaticale devraient apparaître comme la conséquence inévitable d'un état de fait irrémédiable. (De par son aspect idéographique, l'orthographe d'usage pose moins de problèmes : son acquisition ne se substitue pas à un système préalablement acquis, inexistant en l'occurrence.) Plutôt que de stigmatiser les oublieux, en fait nous tous, de les accuser d'illogisme, sans qu'on sache bien ce que la logique a à y voir, mieux vaudrait donc, je crois, les laisser en paix en attendant que les correcteurs automatiques fassent bien leur travail. Je serai satisfait si cet ouvrage, dont le

---

<sup>3</sup> Tel que décrit par Orwell, le *Newspeak* respecte l'orthographe anglaise. C'est dire !



but est précisément de démontrer la différence des systèmes flexionnels oral et orthographique, contribue si peu que ce soit à l'avènement de cette attitude rationnelle.

## 1. Français parlé et français orthographique

« Quand je suis venue en France... j'ai pris des cours de français. C'est une langue difficile à apprendre et illogique. Il y a beaucoup de lettres qu'on ne prononce pas... Il m'est impossible de lire dans cette langue. Le français écrit c'est un autre monde. » Nina, exilée russe (*Le Monde*, 7-8/08/2022, p. 13)

On sait, ou bien Nina nous le rappelle, que le français ne se parle pas comme il s'écrit et inversement. Mais mesure-t-on à sa juste valeur l'écart entre oral et écrit ? Ecart pourtant remarquable et nulle part plus grand qu'en matière de morphologie flexionnelle\* : formation du pluriel et du genre marqué des noms et des adjectifs, déclinaison en nombre, genre et cas des pronoms, conjugaison des verbes, dans tous ces domaines un fossé sépare les deux incarnations de ce qu'on est convenu d'appeler le français. Mais un fossé à la fois évident et toujours dénié.

C'est qu'un écran s'interpose devant ce qui serait une étude réaliste de la flexion\* du français parlé : les descriptions diffusées au-delà du cercle étroit des linguistes professionnels, par le biais de l'écrit traditionnel ou des nombreux sites accessibles sur Internet, s'appuient inmanquablement sur la représentation orthographique de la langue, dont nul n'ignore l'éloignement de la prononciation réelle. Pour pluraliser *chat*, suffixez *-s* ; pour le féminiser, suffixez *-e* (et doublez la consonne finale du singulier) ; la 1<sup>ère</sup> personne du singulier de l'imparfait de l'indicatif de *chanter* se marque d'un suffixe *-ais*, la 3<sup>e</sup> d'un suffixe *-ait* ; *jeter* « prend deux *t* » à telles personnes, un seul à telles autres, etc. Certes, les grammaires soucieuses d'exactitude, telle la *Grande grammaire du français* récemment parue, ne manquent pas de préciser que l'*s* de pluriel ne s'entend le plus souvent pas : « A l'oral, la variation en nombre s'entend rarement : c'est généralement le déterminant (*un, une, des*) qui permet de distinguer le singulier du pluriel » (Melis & Godard 2021). De même la différence purement orthographique entre *chantais* et *chantait* : « La conjugaison est sans doute le domaine où l'orthographe du français s'éloigne le plus de la prononciation. A l'oral, la conjugaison du français se caractérise par un grand nombre de formes ambiguës, ambiguïté qui est largement éliminée à l'écrit par l'utilisation de consonnes finales non prononcées » (Bonami 2021). L'auteur avance prudemment : « utilisation » insinue que, d'un point de vue strictement synchronique, ces « consonnes finales non prononcées » pourraient être un artéfact, un pur procédé orthographique. Mais, qu'on le veuille ou non, comment ne pas accrédi-ter l'idée que le « vrai » français est l'écrit, auquel l'oral serait subordonné, relié à lui par des opérations dont l'effet consiste le plus souvent en une perte d'information ? Ce faisant, on continue de recouvrir le fossé signalé plus haut.

Car l'oral, dès qu'on le considère en oubliant l'écrit, montre des choses bien différentes : /ʃa/ 'chat', à l'instar de tous les noms et adjectifs, à l'exception d'un petit groupe, est invariable en nombre, en sorte que celui-ci n'est marqué en effet que par les déterminants associés, dont la forme phonologique, en dépit de l'orthographe, ne comporte pas davantage le moindre suffixe : /ləʃa/ 'le chat' vs. /leʃa/ 'les chats'<sup>4</sup> ; son féminin est supplétif\*, car aucun phonème ne correspond au <e> orthographique et le /t/ final de /ʃat/ 'chatte' ne se laisse pas analyser comme une marque de féminin ;<sup>5</sup> un suffixe /-ɛ/ marque les trois personnes du

<sup>4</sup> Les transcriptions phonémiques seront explicitées plus loin. La question de la liaison ne sera pas oubliée.

<sup>5</sup> La variété de français parlé ici étudiée est celle d'où les *e* muets (schwas) finaux ont disparu au cours du dix-septième siècle (Brunot 1939 : 190 et suiv.). Je ne dis rien des variétés méridionales où le substrat occitan les a préservés, mais de façon apparemment très variable. La diction des vers classiques, qui les préserve aussi, est sans pertinence pour mon objet.

singulier et la 3<sup>e</sup> du pluriel de l'imparfait de l'indicatif de /ʃâte/ 'chanter' ; /ʒəʒet/ 'je jette' se distingue de /nuʒ(ə)tɔ̃/ 'nous jetons' par l'absence de suffixe et le timbre de la voyelle ; etc. Seuls des linguistes, en petit nombre, ont osé décrire une telle flexion fondée sur ce qui se prononce, en écartant tout à fait la représentation orthographique. On peut craindre que leurs écrits, cités plus loin, n'aient pas eu l'audience qu'ils méritent.

Le but de la présente étude est de contribuer — une fois de plus — à lever l'écran, à découvrir — à nouveau — le fossé qui sépare le français tel qu'il se parle (FP), de ce que je nommerai le « français orthographique » (FO).<sup>6</sup> Mais il faut être clair : comme souligné dans la préface, ce FP que j'entends étudier n'est pas le français dit « populaire » ou « démotique », de même que le FO est autre chose que le français dit « standard » ou, comme on l'ose moins, « correct ». Car le français standard parlé est l'une des variétés de ce que je nomme FP, la variété sociolinguistiquement « haute », et sa flexion est identique à celle des autres variétés plus « basses » et également différente de celle du FO. Si je me suis décidé pour le qualificatif « orthographique », ce n'est pas par souci d'originalité. C'est qu'il convient de prendre au sérieux le caractère écrit et orthographié du FO, qui n'existe que sous cette forme et dont la description grammaticale doit en conséquence se fonder sur cette forme, faite de lettres et non de sons. Pour le dire autrement, les éléments qui expriment la flexion du FO, par exemple *s* de pluriel, sont bien des exposants\*, c.-à-d. des signes associant une forme à une signification grammaticale, en l'occurrence « pluriel », mais ce sont des exposants *graphiques* dont la correspondance avec des exposants phoniques n'est pas donnée d'avance.

Le FP représente ainsi la contrepartie orale, phon(ém)ique du FO — contrepartie, mais non pas équivalent, relation de média à média qui ne vaut pas transposition terme à terme, comme si l'oral n'était que la prononciation de l'écrit et l'écrit que la notation de l'oral, dépendance mutuelle que le français actuel ne cesse de démentir. (Je précise « actuel » pour bien mettre en évidence la perspective synchronique dans laquelle je me place. Evidemment, le français, ou plutôt du français s'est parlé avant de s'écrire, et l'écart entre oral et écrit était bien moins grand au douzième siècle. Ces considérations historiques sont sans pertinence pour la situation en ce début du vingt-et-unième siècle.) C'est pourquoi, pour garder le même exemple, il est inexact d'affirmer, comme le font la plupart des grammaires, que le *s* de pluriel ne se prononce pas. On posera plus justement, me semble-t-il, que le suffixe graphique <*s*> voulant dire « pluriel » du FO est sans contrepartie phonémique en FP. Le résultat est le même, dira-t-on. Sans doute. Mais la formulation que je propose a pour avantage de mettre en relief l'écart qui sépare le FP du FO au niveau de leurs systèmes flexionnels.

Comment du reste ces systèmes pourraient-ils n'être pas disjoints ? N'est-ce pas un fait d'évidence que la flexion que les enfants francophones acquièrent à partir de leur deuxième année, avant toute exposition à l'écrit, ne peut être que celle du FP ? Il faut quand même remettre les choses dans l'ordre : ils n'apprennent pas à suffixer des *s*, mais à ne pas les prononcer ; ils apprennent, d'une part, à associer le même son /ʃa/ au concept CHAT que celui-ci soit instancié par un seul ou plusieurs exemplaires, d'autre part à employer des déterminants différents selon le cas. Plus tard, on leur inculquera (non sans peine) que ce qui se prononce /leʃa/ et dénote une pluralité de félins domestiques s'écrit <les chats>. J'ai l'impression d'enfoncer des portes ouvertes. Mais, ouvertes, le sont-elles vraiment ?

Quant au français « populaire », j'avoue ne pas savoir ce que c'est. Gadet, en conclusion de son étude (1992 : 122), reconnaît qu'il s'agit bien davantage d'un agglomérat de jugements sociolinguistiques que d'une réalité objective. En fait, ce qu'on nomme ainsi

---

<sup>6</sup> Il importe de circonscrire géographiquement l'objet de cette étude. Il n'y sera question que du français parlé dans la moitié nord (pays d'oïl) de l'« Hexagone », qui se trouve être la langue première de l'auteur. Rien sur les français du Québec, des Antilles, de Belgique, de Suisse, d'Afrique ou du Pacifique, langues quant auxquelles l'auteur s'avoue incompetent. Ni sur le français dit « dialectal », car, outre une même incompetence, s'y oppose le fait que gallo, joul, picard, poitevin, wallon, etc. ne sont pas des variétés ou dialectes du français, mais autant de membres du groupe linguistique des langues d'oïl, dont le français fait aussi partie.

n'est le plus souvent qu'une collection de données pittoresques, les plus éloignées possibles du « bon usage », suscitant tantôt réprobation, tantôt approbation, mais qui présupposent dans un cas comme dans l'autre une notion sociopolitique de « peuple » opposé aux « élites » dont les dangereuses liaisons avec les idéologies totalitaires ne devraient plus être à démontrer. C'est pourquoi je préfère le terme « démotique », qui ne partage pas ces connotations et n'implique aucun jugement de valeur, étant entendu que la distinction entre démotique et standard ne recouvre pas celle entre parlé et orthographique.

Il n'est pas question, ce serait absurde, de dénier l'existence d'une variation sociale et socialement évaluée plus ou moins directement associable aux différences de classe qui stratifient la communauté (si c'en est une) linguistique francophone. Le présent travail ne s'en fonde pas moins sur l'hypothèse que cette variation repose sur une uniformité, à savoir une composante flexionnelle commune à toutes les variétés du français parlé. L'idée paraîtra absurde à beaucoup, tant on est habitué à opposer en tout un français normatif à des français plus ou moins déviants par rapport à la norme. Pourtant, aucun francophone, si proche soit son parler d'un standard oralisé, ne produit /le(s)fats/ comme l'eût fait son ancêtre médiéval pour faire référence à une pluralité de chats. On se souvient peut-être de ce discours du général De Gaulle où, ayant énoncé (je cite de mémoire) « Je m'adresse /opœpl/ », il jugea nécessaire de préciser « au pluriel ». Il entendait dire « aux peuples », mais savait bien que cela ne s'entend pas. On le sait, mais on n'en tire pas les conséquences. La distinction FP / FO dans le domaine flexionnel est donc bien indépendante de l'opposition entre français standard et français démotique, elle tient à la différence entre incarnation orale et incarnation écrite. Par là, flexion et syntaxe divergent, pour autant que l'existence de plusieurs syntaxes orales plus ou moins éloignées d'une syntaxe « standard », elle-même sans doute diversifiée, ne peut être mise en doute.

Importante et bien repérée, la variation syntaxique contribue en effet à identifier des variétés, à opposer par exemple une variété normée qui dit « Qu'en conclut-on ? » et « le livre dont j'ai besoin » (oral et écrit) à une moins normée qui dit « On en conclut quoi ? » et « le livre que j'ai besoin » (oral seulement ou écrit l'imitant). En comparaison, la flexion ne varie guère que s'agissant de formes irrégulières et peu fréquentes : qui ne dira jamais « L'économie fleurissait dans les années 60 », alors qu'il faudrait dire qu'elle « florissait » ? Qui n'a jamais hésité sur la conjugaison de *vêtir* : « Vêtissez-vous » ? « Vêtez-vous » ? Relèvent de ces vacillations *ils s'enquérissent* pour *ils s'enquière*, *il acquiera* pour *il acquit* (Gadet 1992 : 52), deux verbes de rare usage et, le second, à un temps pratiquement exclu de l'oral, le passé simple. Les autres exemples de la même autrice, *vous boivez*, *que je save*, etc. me laissent sceptique quant à leur fréquence effective ; j'y vois plutôt des stéréotypes\*, formes rarement entendues, fruits de l'erreur plus que de l'ignorance, mais montées en épingle par des commentateurs aux intentions pas toujours innocentes.

Nulle variation, en revanche, que je sache, quant à l'imparfait d'un verbe fréquent comme *finir* ; aucun *\*finait* comme si le verbe relevait du 3<sup>e</sup> groupe (cf. *sortir / sortait*). D'une façon générale, les verbes « irréguliers » fréquents ne sont jamais régularisés — corrélation entre fréquence et préservation de l'irrégularité de longtemps reconnue par la linguistique historique — et l'on continue à dire *j'ai pris*, mais *j'ai tendu*, quoique les deux verbes partagent une même forme d'infinitif, *prendre* et *tendre* — pourquoi pas *\*j'ai perdu* ? Et si l'on entend assez souvent, mais peut-être pas autant qu'on veut bien le dire, *les journals* ou *Ils sont originals*, c'est que l'invariabilité en nombre des noms et des adjectifs constitue un signe distinctif fondamental de la flexion du français parlé (cf. chapitre 6).

C'est pourquoi la flexion ne remplit pas ou guère de fonction sociolinguistique, cela pour deux raisons. D'une part, sa relative uniformité n'offre guère de brèches aux jugements de valeur. Deuxièmement, de toutes les composantes du langage, la flexion est la plus « dure », la plus purement grammaticale, la moins expressive. Elle n'offre nul espace de jeu

aux points de vue et aux émotions comme le font la phonologie — en particulier dans sa composante prosodique — et la syntaxe. Son système est inaccessible aux jugements épilinguistiques. Tout enfant passé par l'école apprend que *le livre que j'ai besoin* est « incorrect » — ce qui ne l'empêche pas de le dire, mais peut l'amener à se corriger quand les circonstances l'exigent. En revanche, on ne lui enjoint jamais de ne pas prononcer les *s* de pluriel qu'on l'oblige à écrire — quel besoin de le faire ? Pourtant la différence entre construction relative en *dont* et en *que* n'est objectivement pas plus fondamentale que celle entre un système où les noms se pluralisent au moyen d'un suffixe et un système où les noms sont invariables en nombre (comme en chinois, la langue la plus parlée au monde)

C'est cette flexion que je me propose de décrire, c.-à-d. la formation du pluriel et du genre marqué des noms et des adjectifs, la déclinaison en nombre, genre et cas des pronoms, la conjugaison des verbes. C'est elle, répétons-le, que les enfants acquièrent avant toute scolarisation, toute exposition à l'orthographe, qui ne l'affecte qu'à la marge, par exemple en les amenant à faire des liaisons qui ne leur viendraient pas spontanément. (Voir chapitre 3 pour la signification de la liaison.) Justifier cette entreprise, c'est ce que je vais tenter de faire dans les deux chapitres suivants.

## 2. Pourquoi décrire la flexion du français parlé ?

L'auteur du présent ouvrage a dépassé les trois quarts de siècle. C'est un âge où la soif de reconnaissance devrait entrer pour peu dans les entreprises. La question n'est donc que légèrement hypocrite. Que la flexion du français parlé soit à la fois profondément différente de celle du français orthographique et commune à toutes ses variétés, démotiques aussi bien que standard, n'appartient pas en effet à ce réservoir de faits tenus pour vrais sans qu'on éprouve le besoin d'en exiger la preuve ou qu'on possède les connaissances suffisantes pour ce faire, par exemple que notre univers a commencé il y a quatorze milliards d'années dans un grand boum silencieux. (Les créationnistes n'y croient pas et c'est leur droit. J'espère seulement qu'ils resteront minoritaires.) Je ne peux donc me dispenser de justifier au possible mon choix d'en faire une réalité à étudier. C'est aussi un âge où l'on aurait bien tort de ne pas profiter de la liberté d'expression par lui accordée. La question n'est donc pas anodine. Les réponses que je compte y apporter devraient provoquer des réactions, tant les questions de langue sont sensibles dans ce pays. Je le souhaite. Sans trop d'illusions toutefois. Caillou dans la mer plutôt que pavé dans la mare. J'espère quand même quelques ronds dans l'eau.

Sans doute un écart inévitable et sûrement nécessaire entre écrit et oral affecte-t-il toutes les langues pourvues d'une tradition d'écriture. Cet écart, on l'attribue communément à des différences en matière de style et de syntaxe : l'écrit tolère des phrases longues et à enchâssements complexes que, pour partie à cause des limites de la mémoire immédiate, l'oral ne saurait produire. Plus profondément, le passage à l'écrit entraîne une « décontextualisation » du contenu qui impose d'autres conceptualisations et mises en forme (cf. Goody 1977/1979). Cet effet est particulier à l'écriture. La hiérarchie des styles n'est en effet pas inconnue, tant s'en faut, des langues sans tradition d'écriture. Il se raconte en Guinée-Bissau, dans le créole portugais qui en est la langue véhiculaire, le conte de l'enfant qui le parlait le mieux. Je cite : « Le créole est aisé à comprendre... mais une fois que tu l'as compris, tu penses bien avant de parler. Ainsi tu peux raconter tout ce que tu veux à quelqu'un, sans qu'il comprenne rien » (Giusti 1981 : 117).<sup>7</sup> Cette opacité du discours, censée en faire la beauté (salut Mallarmé !), l'enfant l'obtient par l'emploi systématique de

---

<sup>7</sup> En VO : *Kriol sabi obi. Ma ora ku bu kaba obi-l, bu ta kuda diritu anti ku bu na papia. Asin bu ta pudi konta algin tudu ke ku bu misti te i ka ta ntindi nada.* Quoique le kriol, pour l'appeler par son nom, s'écrive de plus en plus grâce aux réseaux sociaux, il n'existe pas d'orthographe officielle. La transcription ci-dessus est ce qui s'en rapprocherait le plus.

métaphores telles que « le feu vêtu d'un manteau » pour le feu recouvert de cendres. Cela pour souligner la probable universalité de l'écart entre l'idiome quotidien et la littérature, qu'elle soit écrite ou orale, bonne ou mauvaise. Mais cet écart stylistique ne se confond pas avec les effets spécifiques de l'écrit, qui se font sentir indépendamment de toute considération esthétique. Parmi eux, d'une particulière importance pour ce qui nous concerne, il y a le fait que la langue écrite ne peut faire autrement que de figer la variation omniprésente dans la langue parlée, sociale aussi bien que dialectale (cf. Cerquiglini 2007, chapitre 7).

Il est donc probable que l'émergence d'un français écrit relativement fixe et normé coïncide avec les premiers documents notant une langue dont la manifestation orale doit sans doute être qualifiée de gallo-romane, mais qui devient (proto-)française du fait précisément d'être couchée sur parchemin. On pense en particulier aux *Serments de Strasbourg* (14 février 842) tels que transcrits par Nithard quelques mois plus tard, dont le caractère transdialectal (ou supradialectal) de la partie romane n'est plus à démontrer (cf. Cerquiglini 1991). Tant que l'écrit français occupera une position subalterne par rapport au latin, qui demeure la langue écrite et parlée de l'administration et de la science au sens large, incluant la philosophie et la théologie, sa graphie continuera d'évoluer au rythme des changements phonétiques qui affectent le gallo-roman septentrional (la langue d'oïl en devenir) du neuvième au treizième siècle — le latin perce encore sous celle de Nithard — mais sans se départir de son caractère transdialectal, qui lui permet d'être employée, moyennant quelques variations, dans tout le pays d'oïl par des auteurs de tous terroirs qui, déjà, ne parlaient pas tout à fait comme ils écrivaient et inversement. Quant à l'expansion à l'oral de ce désormais ancien français, qui doit sûrement beaucoup à un parler parisien dialectalement hétérogène, mais guère à un hypothétique et sans doute inexistant, en tout cas inattesté, dialecte « francien » particulier à l'Île-de-France (cf. Cerquiglini 1991, 2007), elle est due évidemment à des facteurs politiques et économiques bien connus.

Jusqu'au quinzième siècle on peut considérer que les correspondances entre écrit et oral demeurent assez transparentes. C'est que les changements phonétiques et morphosyntaxiques, pour beaucoup à l'œuvre depuis plusieurs siècles, qui ont finalement provoqué l'éclatement d'un latin vulgaire encore relativement unifié en une multitude de parlers romans sont accomplis aux alentours de l'an 600, une centaine d'années après la déposition du dernier empereur romain d'Occident (4 septembre 476). Sans doute ces changements, phonétiques surtout, furent-ils plus extrêmes en gallo-roman septentrional. On invoque généralement l'effet d'un contact plus long et plus intime avec le germanique. Long, car les Francs germanophones s'installent en Gaule du Nord romanophone dès le cinquième siècle ; intime, car contact veut dire qu'au cours du temps, ils et elles — on dit « les Francs », mais il y avait aussi des Franques — sont devenus d'abord bilingues franko-roman, puis unilingues en un gallo-roman dès lors modifié par les habitudes linguistiques de ces nouveaux locuteurs, en particulier la propension à faire porter l'accent tonique par la première syllabe du mot, d'où réduction et chute des voyelles post-toniques. Que cela ait joué un rôle, c'est possible. Mais décisif ? Marcel Cohen se montre sceptique, qui écrit : « Les grands effets de l'accentuation (disparition des voyelles après et avant l'accent) ont parfois été attribués à l'influence germanique, plus particulièrement francique, bas-allemande. De même certaines diphtongaisons. Mais l'ensemble des altérations n'est pas de type germanique » (Cohen 1967 : 109). Je pense qu'il a raison. La chute des post-toniques s'est produite dès le latin vulgaire où l'on prononçait /'oklu/ ce qui s'écrivait encore *oculum* 'œil' — cf. italien *occhio*, portugais *olho*, castillan *ojo*, roumain *ochi*. Le gallo-romain du Nord va certes plus loin en éliminant aussi la voyelle finale ou bien en la réduisant à schwa lorsqu'il s'agit d'un /a/ marque de féminin (cf. *fille*, anciennement /'fiʎə/, comparé à italien *figlia*, castillan *hija*). Mais le premier phénomène touche aussi le gallo-romain du Sud où l'influence germanique semble avoir été négligeable: cf. occitan *uelh*, catalan *ull*, ; et le second s'observe, avec des

résultats divers, en occitan, catalan, portugais et roumain. Les nasalizations, si typiques du français, ne sauraient être imputées au germanique. Bref, vu l'obscurité persistante de l'autre source d'influences possibles, les parlers celtiques réunis sous le nom de gaulois, éteints dès le quatrième siècle, la conclusion la plus sage est sans doute que, oui, le gallo-roman septentrional a connu un ensemble d'évolutions qui le distinguent nettement des autres langues romanes, mais on n'en connaît pas vraiment les causes, sûrement multiples. Surtout, ces évolutions n'ont pas entraîné de mutation typologique. En matière de flexion nominale et verbale, l'ancien français tel que nous le restituent les *Serments*, la *Séquence de Sainte Eulalie* (880), la *Chanson de Roland* (onzième siècle), etc. est une langue romane, et même conservatrice puisque la seule, avec l'occitan, à présenter une déclinaison à deux cas, sujet et objet (régime).<sup>8</sup>

La mutation qui va conduire à ce que Raymond Queneau se plaisait à nommer le « néo-français », c'est dans les bouleversements qui, aux quinzième et seizième siècles, affectent la variété de la langue d'oïl en passe de devenir dominante qu'il faut en chercher les prémices. C'est alors que la linguistique historique trace la frontière entre l'ancien et le moyen français. Pour ce qui est de la flexion, la principale catastrophe alors survenue consiste en l'amuïssement des consonnes finales, que l'ancien français prononçait, qu'elles aient ou non valeur grammaticale. J'ai fait allusion à son effet sur le marquage de la pluralité, toujours lisible, mais devenu le plus souvent inaudible (voir chapitre 6). En flexion verbale, cette même chute des finales entraîne, entre autres, le syncrétisme des trois personnes du singulier et de la troisième du pluriel à l'imparfait de l'indicatif : (*je/tu*) *chantais*, (*elle*) *chantait*, (*elles*) *chantaient* = /ʃãtɛ/. Elles n'en restent pas moins imperturbablement écrites. D'où un écart entre langues écrite et orale qui ne s'en tient pas seulement au niveau du style et de la syntaxe, mais touche aussi le niveau plus essentiel, plus intime de la morphologie flexionnelle, en sorte que la question de l'unité de la langue se pose sérieusement.

Quant aux causes de cette évolution accélérée — on pense à la théorie évolutionniste des équilibres ponctués, applicable à l'évolution des langues (cf. Gould 2006 ; Atkinson et al. 2008) — j'ai bien peur que nous ne soyons contraints à la même conclusion décevante que précédemment. Evidemment, tout ne s'est pas fait du jour au lendemain. La ponctuation du seizième siècle, qui marque la rupture entre ancien français et français moderne, représente le précipité d'un ensemble de changements en cours commencés de longtemps. L'ancien et le nouveau ont coexisté pendant des décennies. Que le seizième siècle soit aussi celui des « Grandes Découvertes » et des guerres de religion, un siècle agité en somme, n'est peut-être pas sans signification, mais le lien causal n'a rien d'évident. Est-il pertinent de faire remarquer que la chute des consonnes finales n'est pas un phénomène limité à la langue d'oïl ? L'occitan provençal et auvergnat le connaissent, avec des conséquences semblables (cf. chapitre 6). L'amuïssement des /s/ de pluriel a touché aussi l'espagnol andalou et certaines variétés d'espagnol latino-américain (cf. Penny 2002 : 106-108). A la même époque, l'anglais a connu une révolution de son système vocalique (*Great Vowel Shift*) qui a modifié profondément l'image sonore de la langue (cf. Chevillet 1994). Il n'existe pas, à ma connaissance, de théorie générale capable d'expliquer ces changements qui affectent

---

<sup>8</sup> On a évoqué l'influence du germanique pour expliquer ce conservatisme isolé. L'hypothèse se soutient de ce que la déclinaison s'est conservée plus longtemps à l'extrême nord du domaine gallo-roman, par exemple en wallon, soit là où le contact avec le germanique a perduré. Mais elle se heurte à l'existence, plus éphémère il est vrai, de cette même déclinaison en occitan, loin de toute influence germanique significative. Une autre hypothèse est que la déclinaison bicasuelle était plus répandue à l'aube des langues romanes, mais que, si seul le gallo-roman la manifeste, c'est à cause de la précocité de son écriture (trois siècles avant l'espagnol, par exemple) qui a pu contribuer à la consolider — précocité elle-même liée, qui sait, à un plus grand écart entre latin et « vulgaire » (cf. Cerquiglini 1991). Un signe peut-être plus assuré de l'influence germanique est le fait qu'en ancien français, comme dans les langues germaniques (sauf l'anglais), le verbe occupe le plus souvent la deuxième place dans la proposition, que le constituant précédent soit ou non le sujet.

simultanément, par des processus semble-t-il indépendants, mais sans doute secrètement reliés, des langues voisines, de près ou de loin apparentées, voire pas du tout.

Aux facteurs historiques sont venus s'ajouter des facteurs sociologiques bien connus : d'une part, l'inertie de l'écrit, toujours en retard sur le changement linguistique, si tant est qu'il le prenne en compte ; d'autre part, la sacralité de l'orthographe française, à peu près fixée depuis le début du dix-neuvième siècle, mais reflétant pour la flexion un état de langue antérieur au seizième, d'où des difficultés arbitraires qui en font un outil de distinction (au sens de Pierre Bourdieu) et, par là, l'un des biens les plus précieux — du moins jusqu'à une date récente — du capital culturel des francophones. Histoire et sociologie conspirent ainsi à masquer la « vraie » flexion du français parlé. Ce complot sans comploteurs, l'orthographe le trahit. Il convient donc de s'y attarder encore un peu et d'en exposer la particularité en la comparant à celle des autres langues européennes, seule comparaison qui fasse sens. (Pour une approche générale de l'orthographe française, voir Beaulieu 1927 ; Fayol & Jaffré 1992 et les autres articles contenus dans cette livraison de *Langue française* ; Riegel et al. 1996, chap. 3 ; Cerquiglini 2004 ; voir aussi les réflexions de Saussure dans le *Cours de linguistique générale*).

Les langues d'Europe qui s'écrivent le font en alphabet latin ou grec ou bien en un alphabet dérivé ou inspiré de ce dernier : cyrillique, arménien, géorgien.<sup>9</sup> Je ne vois que deux exceptions, le yiddish et le judéo-espagnol qui utilisent un alphabet hébraïque adapté. (Mais le second recourt désormais le plus souvent à l'alphabet latin.) Dans presque toutes la relation écrit-oral est assez directe : un ensemble de correspondances régulières permet sans trop d'erreurs de déterminer la prononciation à partir de la forme graphique et inversement.<sup>10</sup> Hormis le français, deux langues semblent s'écarter de ce relatif idéal : l'anglais et l'irlandais.

L'orthographe du premier est notoirement opaque, en particulier s'agissant de la partie latine du vocabulaire. La jeune anglaise qui a entendu le mot /'vi:kəl/ devra retenir qu'il s'écrit *vehicle* 'véhicule' et non, par exemple, \*<veekle>. A l'inverse, la jeune française qui rencontre le mot *vehicle* dans un texte devra apprendre à le prononcer /'vi:kəl/ et non, comme je l'ai entendu, \*/vi'haikəl/, qui n'a rien d'absurde.<sup>11</sup> On connaît la facétie de George Bernard Shaw (1856-1950) qui voulait que *fish* /fɪʃ/ 'poisson' s'écrivît <ghuti>, <gh> comme dans *enough* 'assez' (/ɪ'nʌf/), <u> comme dans *busy* 'occupé' (/ˈbɪzi/), <ti> comme dans les mots terminés en *-tion* /-ʃən/. D'une façon générale, il existe peu de correspondances régulières, dépourvues d'exceptions, entre écrit et oral. L'absence de notation de l'accent tonique, si décisif pour la prononciation des voyelles, est un grave déficit, davantage sans doute pour les étrangers que pour les natifs.

Mais opaque, seule l'est l'orthographe d'usage. L'orthographe grammaticale est, quant à elle, transparente. Tous les morphes flexionnels — il est vrai que l'anglais n'en est pas riche — qui se prononcent s'écrivent et tous ceux qui s'écrivent se prononcent. Comment ils se prononcent varie : cf. *forks* /fɔ:ks/ 'fourchettes', *knives* /nʌɪvz/ 'couteaux', *houses* /haʊsɪz/ 'maisons', *walked* /wɔ:kt/ 'marchais', *decided* /dɪ'saɪɪd/ 'décidais' ; mais c'est du fait d'ajustements phonétiques tout à fait réguliers. Aucune rupture, donc, entre la flexion de l'anglais écrit et celle de l'anglais oral. (Je parle ici, cela va de soi, de l'anglais standard britannique et américain. Il en va bien sûr autrement des anglais « dialectaux ».)

Davantage qu'opaque l'orthographe irlandaise est exceptionnellement compliquée. Elle comporte des correspondances régulières, mais qui n'ont rien d'évident : cf. *abhainn*

<sup>9</sup> Faut-il rappeler que l'Europe abrite de nombreuses langues sans écriture officielle, en particulier celles qu'on qualifie de « dialectes » ? Et aussi que les langues officielles des trois Amériques et de nombreux pays d'Afrique sont des langues européennes ? Les premières ne nous concernent pas, les secondes, si.

<sup>10</sup> En élargissant un peu le domaine géographique, on voit que c'est aussi vrai de l'arabe standard et de l'hébreu moderne. Quant au rapport entre arabe standard et « dialectes », je dirai plus loin ce que j'en pense.

<sup>11</sup> /'viəkəl/ et /'vihikəl/ sont des prononciations admises en anglais américain. Placer correctement l'accent tonique est le problème.

‘rivière’ prononcé /avɾ/, *leabhar* ‘livre’ prononcé /laʊr/, etc. (cf. Ó Siadhail 1980).<sup>12</sup> Reste qu’elle présente la même qualité que l’anglaise de ne pas masquer la flexion, ici fort complexe. Les mutations consonantiques, typiques des langues celtiques, sont soigneusement notées, par exemple l’éclipse dans *cloch* /klox/ ‘pierre’ vs. *ar an gcloch* /əɾə'glox/ ‘sur la pierre’ — il suffit de savoir que, dans le digramme <gc>, seul <g> = /g/, forme éclipse de <c> = /k/, se prononce. Toutes les désinences audibles de la flexion verbale sont graphiquement représentées.

Voilà bien en effet où se situe le problème spécifique de la relation oral-écrit du français. On a coutume de mettre en avant les caprices de l’orthographe d’usage qui font les belles soirées des concours — ainsi le fameux *cuisseau* (de veau) qu’on ne saurait confondre avec le *cuissot* (de chevreuil), ou encore les *choux* vendu dix *sous*. Amusettes, que chacun revendiquera ou déplorera selon son humeur, sa couleur politique, etc. Le vrai problème est que, différente en cela des orthographes des autres langues d’Europe, l’orthographe française note une langue que personne ne parle — puisque personne ne dit \*/lesʃats/ — et que, pour le dire autrement, les francophones parlent une langue qui ne s’écrit pas, puisqu’on n’écrit pas, comme on le pourrait, \*<lé cha> qui transcrirait exactement la flexion orale (voir chapitre 6).

La situation du français s’en voit rendue tout à fait particulière. Prenons, par exemple, le cas bien connu de l’arabe écrit (classique, standard) face aux divers arabes « dialectaux » (algérien, égyptien, marocain, palestinien, etc.). En matière de flexion nominale et verbale, l’écart entre le premier et les seconds est certes plus grand que celui entre FO et FP, quoique, à bien y regarder, guère plus. Mais le monde arabophone vit une diglossie telle que l’existence des dialectes est reconnue, leurs particularités acceptées et leurs domaines d’usage légitime assez bien délimités vis-à-vis de ceux de l’arabe standard.<sup>13</sup> On n’a pas coutume, en revanche, et cela n’aurait pas grand sens, de considérer le FP comme un dialecte, mais comme... quoi au juste ? Non content de manquer d’un terme technique pour le définir, on ne sait même pas le représenter. Car, on ne le dira jamais assez, le FP ne se limite pas au français démotique, puisqu’il partage sa flexion avec le normatif oral, son registre « haut », lequel est lui-même autre chose que le FO. Qu’il en soit syntaxiquement plus proche que les autres registres (« Qu’en conclut-on ? » plutôt que « On en conclut quoi ? »), c’est un fait connu et amplement commenté, comme est notoire l’existence d’écrits qui tâchent de reproduire la syntaxe des registres les moins normatifs du FP, voire (généralement très mal) leur phonologie. Mais peu d’auteurs, et c’est remarquable, se sont aventurés à transcrire la flexion parlée. Raymond Queneau et ses énoncés zaziques (« aboujplu » pour « elle ne bouge plus ») est une exception. Bref, la francophonie est non moins diglossique que l’arabophonie, mais on ne veut pas le savoir (cf. Zribi-Hertz 2013).

Mais « diglossie » est-il bien le terme adéquat ? Revenons sur le cas arabe. L’arabe classique, fondement du standard, possède une flexion nominale à trois cas\* (nominatif\*, accusatif\*, oblique\*) qui a disparu des dialectes. Or, le standard se parle, des situations l’exigent, et qui l’emploie peut prononcer les désinences\* exprimant ces flexions, selon certaines règles explicites. Si l’on désire un autre exemple, on prendra le tchèque, où la différence entre langue normative et langue familière est importante, bien connue et touche la flexion nominale et verbale : cf. normatif *mladý člověk* /'mladi\_ʃlovʲɛk/ vs. familier *mladej člověk* /'mladɛj\_ʃlovʲɛk/ ‘jeune homme’ (nominatif singulier) ; normatif *mohl* /'mohl/ vs. familier *moh* /mox/ ‘(il) a pu’. Mais de même, et sans doute encore plus que l’arabe standard, le tchèque normatif se parle, dans toute situation où il convient de surveiller son langage.

Il en va tout autrement en FP ordinaire : le <s> final de FO *chats*, le <t> final de *elle chantait*, etc. ne sont jamais prononcés hors de certains contextes très limités, liaison et

<sup>12</sup> Le nom de cet auteur, poète et grammairien, se prononce à peu près /o:'ʃi:ʌ/.

<sup>13</sup> Les différences qui les séparent et qui rendent l’intercompréhension souvent difficile font que l’usage tend à s’instaurer de parler de « langues arabes » plutôt que d’« arabes dialectaux ».



inversion. Quant au premier, même le registre haut l'évite à partir d'un nom pluriel comme *chats*, sauf diction poétique ou hypercorrection\* comme dans *des chats égyptiens* prononcé /defæzeziptsjɛ̃/ au lieu du plus courant /defæzeziptsjɛ̃/ et *ils chantaient en cœur* /ilʃætətākœʁ/ au lieu de /ilʃætətākœʁ/. Et on verra au chapitre 6 que le /z/ obligatoire de /lezami/ n'a qu'un rapport indirect au premier <s> graphique de *les amis*, tandis que le /t/ également imposé de *chantait-elle* ? /ʃætətɛl/ n'est plus, et depuis fort longtemps, une désinence personnelle. La situation du français apparaît ainsi encore plus singulière que dit plus haut. On n'a pas affaire comme dans les diglossies classiques à deux langues ou deux variétés également effables, quoique dans des environnements distincts, mais à deux manifestations dont l'une n'a pratiquement d'existence qu'écrite s'agissant d'un domaine fondamental de sa grammaire. « Disystémie » plutôt que diglossie ? On évitera les néologismes superflus... mais on ne s'étonnera plus que nos textes, même écrits par des bac + n, fourmillent de <s> oubliés, de désaccords, d'*Omar m'a tuer*, etc.

Bref, s'extraire de cette confusion demande, me semble-t-il, de procéder comme le ferait une linguiste sénégalaise débarquant à Nogent-sur-Marne (ou à Charleville, ou à Bourges) pour y étudier la langue locale non écrite et non décrite, dans une uchronie où les colonisations se seraient faites à rebours. Et si, de ce fait, le FP prend l'allure d'une langue « exotique », tant mieux. C'est que j'aurai réussi à faire voir combien les représentations qu'on s'en fait communément sont modelées par la grammaire du français écrit, non seulement traditionnelle, mais aussi telle que la transmettent nombre d'études à ambition scientifique. Du reste, exotique, la flexion du FP l'est vraiment si on la compare à celle des autres langues romanes, alors que la flexion du FO exhibe encore fièrement sa latinité.

Mais cette ambition est-elle réaliste ? A-t-on le droit d'oublier que nous vivons dans une société lettrée à 93% (selon l'INSEE), et que cela ne peut pas ne pas influencer les savoirs et les usages ? Ne voit-on pas qu'en cas de doute sur le nombre singulier ou pluriel d'un syntagme nominal bien des locuteurs choisissent le recours de s'en référer à l'écrit et de préciser ou bien « au pluriel », comme le fit, on l'a vu, le général De Gaulle dans un discours où il proclamait s'adresser « aux peuples » (/opœpl/), ou bien « avec un s » ?<sup>14</sup> Il me semble que oui, pour la simple raison que, comme dit plus haut, l'essentiel, voire la quasi totalité de la flexion est acquise avant toute scolarisation. La flexion du FO enseignée à l'école se superpose à cette compétence naturelle, mais elle ne saurait s'y substituer. Tout au plus, la relation FP-FO étant de type diglossique, moyennant la particularité signalée plus haut, va-t-elle la parasiter, avec des effets variables, parfois utiles à la communication.

Ici, il me faut à nouveau être clair. Il est évident que, pour des raisons politiques et sociologiques qu'il n'est pas besoin de détailler, l'école se doit d'enseigner le FO, en particulier l'orthographe grammaticale qui, pour être complexe, se révèle plutôt bien adaptée au code écrit. On est seul face au texte, ce qu'il ne consigne pas n'est pas récupérable ; l'échange oral *in praesentia* tolère autrement plus d'ambiguïtés et de lacunes dans l'expression. Quant à l'orthographe dite d'usage, les diverses propositions de simplification ou bien, trop timides, ne font guère que déplacer les difficultés — en quoi *ambiguë* est-il plus simple qu'*ambiguë* ? — ou bien, trop radicales (*ābigy*), ne peuvent que susciter le refus d'une telle altération de l'image de la langue écrite. L'orthographe du français a tout d'un monument national. On l'entretient, on peut la réformer ponctuellement, on ne la reconstruit pas. Sensibiliser les élèves à la variation sociolinguistique est également nécessaire. On peut le déplorer, mais il est des circonstances où « ce dont j'ai besoin » plutôt que « ce que j'ai besoin » peut faire la différence entre un emploi et le chômage.

<sup>14</sup> Voilà qui rappelle l'usage chinois de tracer virtuellement du doigt le caractère dans la paume de la main en cas de doute sur l'identité du mot oral. L'analogie entre l'orthographe française et les écritures idéogrammatiques a souvent été soulignée.

Mais ces considérations conservatrices, voire réactionnaires, ne sont pas mon dernier mot. Car si l'école a pour tâche d'enseigner à ses élèves un français aussi proche que possible de la norme édictée par les textes officiels et la plupart des grammaires, peut-être y réussirait-elle mieux si elle se délivrait de l'idée que c'est là la seule bonne langue, qu'il faut acquérir comme si l'on ne disposait d'aucun vrai langage tant qu'on n'a pas celui-là, plutôt qu'une langue qu'il convient de savoir lire, écrire et, au possible, employer dans les situations qui l'exigent, un peu comme on met une cravate ou un tailleur. Ne vaudrait-il pas mieux initier les élèves aux différences entre leur parler natif pré- et extra-scolaire, parfaitement expressif et fonctionnel, auquel le cours de « français » ne devrait rien avoir à reprocher, et le code normatif qu'elles et ils auront à maîtriser dans leur vie active, les choses étant ce qu'elles sont ?<sup>15</sup> Autrement dit, leur enseigner le français normatif un peu — je dis bien « un peu », car la situation n'est pas identique — comme on enseigne une langue étrangère indispensable, le latin pour qui aspirait à faire des études universitaires au Moyen Âge, l'anglais de nos jours. L'enseignement des langues régionales aiderait aussi. Par exemple, l'occitan languedocien est plus proche du FO que du FP en matière de flexion nominale et verbale. Mais pour cela, encore faut-il avoir une connaissance explicite de la différence objective qui sépare les deux systèmes, en particulier s'agissant de la flexion.

Cette même connaissance aiderait aussi sans doute à traiter rationnellement et sans passion un problème qui empoisonne la vie quotidienne de beaucoup : les fautes d'orthographe qui, pour la plupart, ne font que trahir un point de contact entre FO et FP : écrire *cinq livre* sans <s>, qu'est-ce sinon infiltrer la flexion du second dans le premier ? Mais le comprendre, c'est comprendre qu'il n'y a pas là de « faute », encore moins de déficit intellectuel.

En 1982 le grand linguiste américain William Labov fut appelé à témoigner comme expert dans un procès que des familles afro-américaines avaient intenté à l'Etat du Michigan. Ces parents se plaignaient de ce que leurs enfants étaient discriminés en classe à cause de leur parler où les professeurs ne voyaient qu'anglais déformé, que tournures « illogiques ». Labov démontra qu'en réalité ces enfants s'exprimaient en une variété d'anglais propre à leur communauté, l'*Afro-American Vernacular English* (AAVE), cohérente et d'égale valeur à toutes les autres variétés d'anglais. Le juge donna raison aux plaignants et enjoignit aux écoles de l'Etat de ne plus chercher à extirper ce vernaculaire de l'esprit des élèves afro-américains, mais de désormais leur enseigner l'anglais standard comme une autre variété utile à connaître. Ce verdict fut-il suivi d'effets durables ? Je l'ignore. Pouvait-il l'être quand les Afro-Américains continuent à souffrir de discriminations sur bien d'autres choses que leur parler ? Il est permis d'en douter. Reste l'exemple, qu'il faudrait suivre.

Ou plutôt qu'il *faudra* suivre. J'ai qualifié l'orthographe française de bien culturel. Est-ce encore vrai ? Ou bien n'assistons-nous pas à un retour à ce qui fut la situation ordinaire jusqu'aux normalisations du dix-neuvième siècle, quand la notion de faute n'existait guère, et l'on n'était pas jugé sur des <s> manquants ? Vision aristocratique, dira-t-on, propre à un monde où la valeur venait de la naissance, non du mérite acquis. Sans doute. Mais vision qui pourrait bien retrouver sa légitimité dans notre monde où le contrôle d'un écrit démultiplié par la technique devient toujours plus impraticable par les moyens traditionnels.

Au fond, si les correcteurs de nos ordinateurs et téléphones intelligents étaient capables d'intervenir non seulement sur l'orthographe d'usage, ce qu'ils font déjà plutôt bien, mais aussi, automatiquement et sans erreur, sur l'orthographe grammaticale, l'orthographe comme « science » rejoindrait l'héraldique au rayon des savoirs non dénués d'intérêt, mais de peu de valeur sociale. Vu la complexité théorique et technique du problème, il est toutefois probable que les cours d'orthographe aient encore de beaux jours devant eux. Mieux vaut

---

<sup>15</sup> Je laisse évidemment de côté la créativité littéraire qui échappe à tout enseignement et ne touche qu'une faible partie de la population.

faire en sorte, si cela se peut, que la « faute » cesse d'être un moyen de distinction sociale. Il se pourrait du reste que la chose soit déjà et d'elle-même en bonne voie.

### 3. Sauver l'orthographe ?

« L'ortograf est une mandarine »

Écrit sur les murs en mai 1968

Opposer flexions du FP et flexions du FO oblige à démontrer l'inaptitude de la graphie traditionnelle à noter les premières. Démonstration qui ne va pas de soi, tant il est vrai que plusieurs travaux linguistiques ont entrepris de « sauver l'orthographe », c.-à-d. d'élaborer des analyses qui, sans que c'en soit le but premier, ont pour effet de soutenir que les représentations orthographiques, éloignées qu'elles sont de la prononciation, n'en donnent pas moins une représentation exacte de la flexion du français, oral aussi bien qu'écrit. Pour ce faire, on s'est beaucoup appuyé sur deux phénomènes : la formation du féminin et du pluriel des noms et des adjectifs (cf. en particulier Dell 1973 ; Schane 1973 ; Huot 2005). Il en sera question en détail aux chapitres 6 et 7. Je n'en dirai ici que le nécessaire pour mon propos.

Soit l'expression du genre grammatical des adjectifs, par exemple de paires masculin / féminin telles que /gʁo/ <gros> vs. /gʁos/ <grosse> et /lɔ̃/ <long> vs. /lɔ̃g/ <longue>.<sup>16</sup> A s'en tenir aux réalisations notées en alphabet phonétique, le féminin se distingue du masculin par la présence d'une consonne finale, absente du masculin, de nature variable et le plus souvent imprévisible comme le confirmerait une liste plus fournie. Conclusion inacceptable pour beaucoup, soit qu'elle leur paraisse trop étrangère à l'image commune du français, soit — et c'est la réaction de nombreux linguistes — à cause de son absence de généralité due précisément au caractère imprévisible de la consonne finale additionnelle associée au féminin, qui a pour conséquence que les paires de cette sorte doivent être mémorisées une à une, sans possibilité de poser une règle générale — ce qui ne peut que déplaire à tout esprit scientifique.

Or, l'orthographe, on le voit sur les exemples, note ces consonnes dès le masculin, d'où l'hypothèse que c'est elle qui dit la vérité des formes, et que /gʁos/ et /lɔ̃g/ représentent les formes « sous-jacentes » de ces adjectifs, c.-à-d. les formes sous lesquelles on les suppose rangées dans le lexique mental des francophones, si l'on se soucie de réalisme cognitif ; ou bien simplement les formes qui permettent de poser une règle, si l'on ne s'en soucie pas ou guère. Au masculin, genre non marqué, les consonnes /s/ et /g/ se trouvent en finale de mot, d'où elles disparaissent par l'effet d'une règle phonologique de troncation affectant de façon générale, mais non sans de nombreuses exceptions, les obstruantes finales (Dell 1973 : 182-186).<sup>17</sup> Au féminin, si l'on admet que le <e> orthographique correspond au suffixe de genre longtemps prononcé /-ə/ (cf. note 2), les consonnes /s/ et /g/ ne sont plus finales « en réalité » et échappent ainsi à la règle de troncation (cf. Dell 1973 : 185-186 ; Huot 2005 : 133).

Cette analyse est insatisfaisante par plusieurs aspects. L'hypothèse du /-ə/ suffixal, désormais inaudible pour la plupart des francophones, y compris là où l'influence des langues d'oc amène à prononcer davantage de *e* dits « muets », est fragile, comme est fragile tout ce dont l'existence ne se soutient que de son utilité heuristique. A l'appui de la réalité de cet élément, Dell (1973 : 186) fait remarquer qu'il se prononce lorsque le mot suivant commence par un « *h* aspiré » empêchant liaison et enchaînement, d'où le contraste <grosse ourse>

<sup>16</sup> Les parenthèses angulaires encadrent les formes orthographiques (voir chapitre suivant).

<sup>17</sup> Pour mémoire, sont obstruantes les consonnes dont l'articulation exige la fermeture totale ou partielle du canal vocal, soit en français /p/, /t/, /k/, /b/, /d/, /g/, /f/, /v/, /s/, /z/, /ʃ/, /ʒ/. Les liquides /l/ et /r/ et les nasales /m/, /n/, /ɲ/ et /ŋ/ ne sont pas obstruantes. La règle possède un fondement historique : en ancien français /gros/ masculin s'opposait à /grosə/ féminin. La chute des consonnes finales en moyen français et l'amuïssement complet des *e* dits « muets » finaux au cours du dix-septième siècle ont mené à la situation actuelle.

/gʁosuks/ vs. <grosse housse> /gʁosœus/. Le contraste est réel, sauf que la réalisation /gʁosus/ ne me paraît pas mériter l’astérisque dont l’accable Dell. Non seulement la prononciation sans hiatus est acceptable dans ce cas, mais elle l’est encore plus, me semble-t-il, quand le mot en *h* n’est pas monosyllabique : cf. <une grosse hallebarde> /yngʁosalbɑrd/, <une grosse huguenote> /yngʁosygnɔt/, etc. S’il y a quelque bizarrerie dans /gʁosus/ ou encore /gʁosaʃ/ (<grosse hache>), /gʁosɛ/ (<grosse haie>), /gʁosi/ (<grosse hie>), etc., elle provient plutôt de la difficulté qu’on peut éprouver à segmenter exactement ces suites phonétiques, en particulier les deux dernières. Cela ne constitue en rien une preuve empirique de la réalité en FP d’un suffixe /-ə/ exprimant le genre grammatical féminin à la manière du suffixe /-a/ des autres langues romanes. Une autre explication est possible : voyelle neutre, schwa (nom technique de /ə/) représente le segment spontanément inséré pour casser les séquences phoniques ressenties comme d’une quelconque façon malcommodes ou anormales : cf. <pneu> prononcé /pœnø/, <immense stade> prononcé /imɑ̃sœ\_stad/, <T-shirt jaune> prononcé /tifœʁtœʒon/ (Dell 1983).<sup>18</sup> Son insertion facultative dans /gʁosœus/ et quasi obligée dans /gʁosœi/ procéderait de cette fonction normalisatrice : l’hiatus ainsi créé rend plus audible qu’on a affaire à deux mots. Mais s’il n’existe pas de suffixe /-ə/ féminin, l’explication par troncation de l’alternance /gʁo/ vs. /gʁos/ ne tient plus.

Que la forme du féminin serve de base de dérivation — cf. <grossiste>, <longueur>, etc. — n’est pas pertinent en l’occurrence. Ici, les consonnes prédésinentielles /s/ et /g/ sont bien suivies d’une voyelle (/i/, /œ/), en sorte que leur réalisation en démontre en effet la réalité, qu’il n’a jamais été question de nier. Mais cela n’en explique pas l’apparition au féminin, s’il est vrai qu’elles n’y sont suivies de rien.

La règle de troncation se révèle en outre n’être pas aussi générale qu’elle le devrait. Il faut en effet distinguer entre, p.ex., /lɔ̃g/, censé y être soumis, et /dɛ̃g/ qui y échappe car épïcène. Certes, l’orthographe indique la différence en écrivant <long> et <dingue>. Mais le <e> final du second n’est pas la marque du féminin, ni de rien sinon de l’audibilité du <g> précédent, dont <u> indique la prononciation /g/ plutôt que /ʒ/. A noter encore les consonnes finales « fixes », toujours prononcées, d’adjectifs dont le féminin n’en est pas moins idiosyncratique : p.ex. *net* /nɛt/ vs. *nette* /nɛt/, de fait épïcène malgré l’orthographe, *grec* /grɛk/ vs. *grecque* /grɛk/ (même remarque), *sec* /sɛk/ vs. *sèche* /sɛʃ/ ; et aussi l’imprévisibilité persistante de la consonne censée révélée par le blocage de la troncation : partant de formes « sous-jacentes » comparables telles que consignées par l’orthographe, comme /rus/ (<roux>) et /ʒalus/ (<jaloux>), pourquoi le féminin de l’un est-il /rus/ (<rousse>), mais /ʒaluz/ (<jalouse>) celui de l’autre ? A moins que la consonne sous-jacente du second ne soit /z/. Mais comment le sait-on autrement qu’en partant du féminin, ce qui rend le raisonnement circulaire ?

Cette interrogation soulève le problème fondamental de la plausibilité cognitive de l’hypothèse de la troncation. Comme il est bien évident que les jeunes francophones acquièrent ces paires masculin-féminin bien avant que l’école ne leur ait donné la moindre notion d’orthographe, tout ce dont on peut être certain, c’est qu’elles/ils (a) entendent les deux formes avec et sans consonne finale ; (b) comprennent qu’elles sont de même sens mais d’extensions différentes ; (c) observent que les sans consonne et les avec consonne forment deux ensembles en distribution complémentaire, en sorte qu’il n’est pas permis de les échanger dans un même contexte. A ces constatations nécessaires et a priori suffisantes, on vient alors ajouter l’hypothèse qu’elles/ils découvrent que la présence de la consonne finale est due à celle d’un élément porteur de la signification « féminin », mais pratiquement jamais audible — car la clause du *h* aspiré est manifestement insuffisante pour lui conférer une réalité empirique. Les voies de cette découverte pré-orthographique restent mystérieuses.

<sup>18</sup> Effectivement prononcée, la voyelle neutre se réalise en FP comme la voyelle antérieure arrondie mi-ouverte que l’API note /œ/.

Il apparaît ainsi que l'hypothèse de la troncation satisfait mal au principe de parcimonie, si du moins on la prend pour une simulation plausible de l'acquisition, puis de la connaissance implicite des paires masculin-féminin du type /gʁo/ vs. /gʁos/. A moins bien sûr que le schwa final ne soit un artéfact formel permettant de poser une règle générale mais n'entraînant nulle conjecture quant à l'état cognitif des sujets parlants. Mais ce n'est pas dans cette optique que l'hypothèse a été formulée.

La question du pluriel nous retiendra moins pour le moment. L'analyse, disons, conservatrice consiste, on l'a vu, à poser que le pluriel des noms et des adjectifs se forme, sauf exception, par la suffixation d'un *-s*, qui s'écrit mais ne se prononce pas, sauf en contexte de liaison qui en révélerait l'existence : cf. *les copains* {le-s\_kopɛ̃-s} /lekopɛ̃/ vs. *les amis* {le-s\_ami-s} /lezami/ (cf. Huot 2005 : 134-136). A cela on peut objecter que la liaison est un phénomène très complexe, trop sans doute pour faire un fondement ontologique fiable. J'y reviendrai au chapitre 6 dans la section consacrée au marquage du nombre.

Le même raisonnement, appliqué à la flexion verbale, amène à dégager une désinence /t/ à la 3<sup>e</sup> personne de l'imparfait, elle aussi révélée par une liaison, obligatoire dans *Chantait-elle ?* /ʃātɛtɛl/, facultative dans *Elles chantaient en chœur* /ɛlʃātɛ(t)ākœʁ/. A l'objection touchant à la fragilité de la liaison en tant que principe d'explication, on peut ajouter ici que cette désinence aurait pour particularité de ne signifier que la personne, à l'exclusion du nombre, puisque <en> qui différencie <chantait> de <chantaient> est purement orthographique depuis des siècles. Or, les désinences qui s'entendent toujours, par exemple le *-ons* /-ɔ̃/ de *chantons*, le *-t* /-t/ de *vous fîtes*, etc. expriment toutes conjointement personne et nombre. Il faudrait donc rendre compte du caractère exceptionnel de ce /t/ censément de 3<sup>e</sup> personne, mais indifférent au nombre.

Tout cela soulève un problème épistémologique que je ne peux qu'énoncer dans les limites de cet ouvrage : peut-on légitimement supposer l'existence de causes si bien cachées qu'elles ne se manifestent jamais que par certains effets qu'elles sont censées produire et, ce faisant, expliquer ? Les supposer, en faire l'hypothèse, oui bien sûr ; mais estimer l'hypothèse confirmée et l'élever au statut de fait établi — par exemple la réalité grammaticale d'un signifiant *-s* porteur du sens « pluriel » ou d'un signifiant *-t* de 3<sup>e</sup> personne du singulier ou du pluriel —, sûrement non, tant qu'il n'est pas acquis que les phénomènes qu'on cherche à expliquer ne peuvent l'être autrement et de façon plus plausible, plus parcimonieuse au sein du domaine où ils se manifestent. On verra que de telles explications sont en effet disponibles. Et j'ajoute, car on n'est jamais assez clair, que, puisque la réalité *orthographique* de <s>, <t>, etc. est indiscutable, il est tout à fait normal d'en faire état dans une description du FO. La question est : est-il légitime d'en faire état dans une description du FP ? Question qui ne peut évidemment être posée si l'on ignore ou dénie la différence systémique entre FO et FP.

#### 4. Comment représenter le français parlé ?

Comme toute langue orale, le FP ne s'étudie bien qu'écrit, c.-à-d. transcrit dans un code qui en fixe aussi précisément que possible la réalisation. On peut le regretter et filer la métaphore du papillon formolé épinglé dans la boîte. Mais c'est la condition pour qu'il y ait une entomologie et une science du langage. Une grammaire n'est pas un film, c'est une photo. Elle encadre la langue, dont les limites sont floues — qu'est-ce qui se dit ? qu'est-ce qui ne se dit pas ? — ; elle l'immobilise, la toujours changeante. L'important est que la photo soit nette et sans retouches ; et qu'on n'en fasse pas une icône sacrée.

Il existe une tradition de transcription du français « populaire » à base d'y (*y vient pour il vient*) et d'apostrophes (*Nini peau d'chien*). Erratique et condescendante, elle n'est évidemment d'aucune utilité. Plus proches de mes besoins sont les transcriptions pseudo-phonologiques de Raymond Queneau : *doukipudonktan*, *aboujplu*, etc. Il me paraît en effet

aller de soi qu'une transcription phonologique systématique utilisant l'alphabet phonétique international (API) s'impose pour une analyse objective de la flexion du FP, délivrée du brouillage orthographique — et pour cela seulement, car il va également de soi — sans doute faut-il le redire tant les sensibilités sont à vif sur ce point — que je n'en préconise nullement l'adoption pour l'usage ordinaire. Couramment utilisé dans les dictionnaires — et même, comme j'ai pu le constater récemment, dans les affiches publicitaires — l'API tel qu'appliqué au français me paraît désormais assez largement connu pour que son emploi ne cause pas de difficulté.

Les mots et les phrases du FP seront systématiquement traduites en FO. Cela peut paraître artificiel. Je l'estime néanmoins utile, voire nécessaire pour faire sentir qu'en matière flexionnelle (et pas seulement) FP et FO représentent deux systèmes typologiquement distincts. Pour le FO, j'utiliserai évidemment l'orthographe traditionnelle, entre apostrophes pour les traductions, ainsi *vaf* 'vache(s)', en italiques quand en citation dans le texte — la différence d'alphabets devrait suffire à éviter toute confusion —, entre parenthèses angulaires (<vache>) quand il me paraîtra utile de faire ressortir le caractère *graphique* des mots du FO.

Comme l'exemple *vaf* 'vache(s)' suffit à le démontrer, la différence typologique a pour effet que la correspondance FP-FO est loin d'être toujours biunivoque et qu'à une forme du FP correspondent souvent plusieurs formes du FO : à l'unique forme FP *vaf*, qui ne marque pas d'elle-même le nombre, correspondent deux graphies FO, <vache> singulier et <vaches> pluriel (voir chapitre 6) ; de même *ŷāt* et <chante>, <chantes>, <chantent> (voir chapitre 9). L'inverse est plus rare, mais il se rencontre : cf. FO <gros ennui> et FP *gro\_ānyī* ou *groz\_ānyī*.

J'aurai recours à deux sortes de transcriptions : lexématique en italiques sans encadrement (p.ex. *vaf* 'vache'), phonémique en romains entre barres obliques (p.ex. /vaf/). S'agissant des mots simples, les deux coïncident généralement, à l'exception des <e> « muets » internes labiales par définition, notés ə (schwa) en transcription lexématique (p.ex. *ātrəpriz* 'entreprise'), /œ/ ou /ø/ en transcription phonémique quand ils se prononcent (/ātœpɾiz/).<sup>19</sup> Quant aux <e> finaux, tel celui de <vache>, je ne les note pas, les considérant comme purement orthographiques et sans correspondant phonologique potentiellement porteur de sens en discours ordinaire — c'est-à-dire hormis la diction traditionnelle du vers classique, par exemple — et cela même pour les locuteurs méridionaux qui réalisent davantage de schwas internes. Cette décision n'est pas, on le verra, sans conséquences morphologiques. Comme le montre aussi l'exemple d'*ātrəpriz* 'entreprise', et afin de tenir compte des variations possibles, les transcriptions lexématiques ne cherchent pas, par principe, à coller à la réalité phonétique, d'où, par exemple, ma notation par *r* de ce qui, pour moi, est la fricative uvulaire voisée /ʁ/, mais peut avoir d'autres réalisations.

Mes transcriptions phonémiques sont du reste elles aussi assez larges. Ainsi, je note toujours les liquides finales /l/ et /r/ précédées d'une consonne obstruante (/p/, /t/, /k/, /b/, /d/, /g/, /f/, /v/), par exemple *kabl* 'cable' et *katr* 'quatre', en fait prononcés /kabl/ ou /kab/, /katr/ ou /kat/. En principe, l'effacement de /l/ et /r/ dans cet environnement n'a lieu, de façon toujours variable et sociolinguistiquement évaluée, que devant une pause ou un mot à initiale consonantique : cf. *pete\_ē=kabl* 'pêter un cable' /peteēkabl/ ou /peteēkab/, *katr\_kabl* 'quatre cables' /katrœkab(l)/ ou /katkab(l)/ (cf. Dell 1983). J'ignore ces phénomènes parce que, outre qu'ils ne touchent pas à la flexion, ils sont de nature phonétique — et fort anciens : Froissart (~1337~1410) fait rimer *troeve* (*trouve*) avec *oeuvre* (*ouvre*), dénonçant malgré l'orthographe une prononciation /œv/ et confirmant du même coup, s'il en était besoin, que la différence entre FO et FP ne date pas d'hier (cf. Zink 1990 : 15). Peu de francophones effacent /l/ et /r/ post-obstruante aussi devant voyelle, voire n'ont d'autre forme que /tab/ et /kat/ dans leur

<sup>19</sup> Je précise que le système phonémique retenu est le mien. Cela ne devrait avoir aucune incidence sur la description morphologique.

lexique mental, et prononcent, à l'occasion ou régulièrement, /katami/ pour *katr ami* 'quatre amis' — mais il y en a (cf. Laks 1980). Le schwa qui se fait entendre dans /katrœkab(l)/ ou /tablœbas)/ 'table basse' n'a pas à être noté, car il résulte d'une épenthèse à peu près obligée pour des raisons articulatoires dès lors qu'on se décide à prononcer la liquide. Il n'apparaît jamais quand celle-ci est absente (\* /katœkab(l)/, \* /tabœbas)/. Je ne tiens pas davantage compte de la palatalisation des occlusives, en particulier /t/ et /k/, devant les voyelles hautes (/i/ et /y/), qui s'observent dans plusieurs variétés, pas seulement au Québec.

Toute chaîne ininterrompue de caractères alphabétiques (p.ex. *vaf*) correspond à un simplex, c.-à-d. à un mot-forme morphologiquement inanalysable. Par convention, les lexèmes que les mots-formes réalisent sont notés en petites capitales et en orthographe traditionnelle, p.ex. VACHE. (Ces notions et celles introduites à la suite seront explicitées au chapitre suivant.) Enrichies d'opérateurs, les transcriptions lexématiques deviennent des transcriptions morphosyntaxiques notant les suites complexes : tiret pour séparer les affixes flexionnels de la base : p.ex. *rymin-e* 'ruminer | ruminé.e.s' ; « = » pour unir les mots clitiques à leur hôte : p.ex. *le=ryminasjǽ* 'les ruminations' ;<sup>20</sup> tiret souscrit entre les mots non clitiques en construction et formant une unité prosodique : p.ex. *le=profǽd\_ryminasjǽ de=vaf* 'les profondes ruminations des vaches' ; puce pour séparer une consonne de liaison de ce qui précède : p.ex. *Dǽn=mwa•z=ǽ* 'Donne-moi-z'en', *M=em•t=el* ? 'M'aime-t-elle ?'. La barre verticale dans les traductions signale une disjonction : *rymin-e* est la forme ou bien de l'infinitif, ou bien du participe passé, masculin ou féminin, singulier ou pluriel, comme noté par les points « inclusifs » dans la traduction en orthographe traditionnelle. La distinction infinitif / participe est active en FP, mais les formes en sont identiques pour les verbes du 1<sup>er</sup> groupe (cf. chapitre 11). Cette identité formelle est source de fréquentes « fautes d'orthographe », comme dans cet avis collé sur la porte d'un restaurant : « Fermer ce soir » (observation personnelle). On n'en conclura pas que, comme le prétendent des esprits chagrins, les incultes ne distinguent plus l'infinitif du participe. Il est bien évident que si la ou le propriétaire avait voulu au contraire rassurer sa clientèle, elle ou il aurait écrit « Ouvert ce soir » et non « Ouvrir ce soir ». Lorsqu'une même unité se réalise de plusieurs façons sans que cela en affecte le sens, je sépare les variantes par un tilde : p.ex. *sǽlyila ~ syila* 'celui-là'.

Mon choix d'une notation aussi économe que possible, entièrement linéaire et sans structure hiérarchisée repose sur deux considérations. D'une part, il m'est apparu, et j'espère qu'on s'en convaincra, qu'elle suffisait amplement à représenter de façon adéquate la flexion du FP. D'autre part, elle est justifiée par la conception de la morphologie flexionnelle adoptée dans cet ouvrage et explicitée au chapitre suivant.

## 5. La morphologie flexionnelle : vue d'ensemble

### 5.1. Introduction

Au long des chapitres qui viennent j'explorerai successivement les flexions nominale, adjectivale, pronominale et verbale du FP qui, que je sache, n'en connaît pas d'autres. Comme promis, je m'efforcerai de rendre l'exposé accessible au plus large lectorat possible. Il me paraît toutefois évident que je n'y parviendrai qu'à condition de lui fournir sur le domaine exploré un ensemble d'informations que ni les cours de français, ni les éventuels cours d'introduction à la linguistique n'ont pu lui donner, car ce n'était pas leur objet. M'excusant auprès de celles et ceux pour qui une telle précaution est inutile, je supposerai donc chez qui me lit une absence de connaissances sur ce qu'est la morphologie flexionnelle, au-delà du fait que les verbes se conjuguent et — pour les *happy* (?) *few* qui ont « fait » du latin, du grec, de

<sup>20</sup> Ce travail étant consacré à la seule flexion, je m'abstiens le plus souvent de noter les frontières dérivationnelles des mots construits tels que *ryminasjǽ* 'ruminations'. Si nécessaire, je le fais au moyen d'un dièse (*rymin#asjǽ*).

l'allemand ou du russe — que les noms se déclinent. Le présent chapitre a pour but, pour prétention de remédier à cette bien innocente ignorance.

Implicite dans cette présentation est l'hypothèse fondamentale, qui ne fait pas l'unanimité comme on le verra, selon laquelle la morphologie constitue une composante autonome de la grammaire, distincte des autres composantes par ses principes et ses procédures. Autonome, mais pas isolée. Les flexions, par exemple les diverses formes de la conjugaison d'un verbe, expriment des significations, ainsi la chronologie de tels événements, elles ont donc des contenus sémantiques. D'autre part, entre flexions et syntaxe il existe à l'évidence des liaisons, des interfaces où s'échangent les informations, par exemple sur le genre grammatical et le nombre pour l'accord entre noms et adjectifs : cf. *lə=fa\_blä* 'le chat blanc' vs. *la=fat\_bläf* 'la chatte blanche'. Les buts de mon étude font que j'aurai peu à dire quant au versant sémantique. Les termes de la grammaire traditionnelle — présent, passé, singulier, pluriel, etc. — me suffisent et devraient être familiers à qui me lira. (Cette familiarité cache évidemment bien des obscurités et complications, mais nous pouvons ici les ignorer sans dommages.) Quant à la syntaxe, quand il me faudra en traiter, j'adopterai pour l'essentiel le modèle et la terminologie de la Grande Grammaire du français (Abeillé & Godard 2021) dont l'approche « surfaciste » — l'analyse doit s'appuyer sur la structure apparente, non sur une quelconque structure « profonde » qui en différerait — se marie au mieux avec la conception ici défendue de la morphologie flexionnelle.

## 5.2. Mots, lexèmes, mots-formes, paradigmes

En linguistique comme en biologie, la morphologie est l'étude (λόγος) de la forme (μορφή) externe et interne (structure), des organismes dans celle-ci, des unités signifiantes dans celle-là. Le terme ordinaire pour « unité signifiante » est alors « mot ». La morphologie linguistique est l'étude de la forme et de la structure des mots. Qu'est-ce qu'un mot ? On croit le savoir. Mais il se pourrait que, comme le temps pour St Augustin, ce soit une de ces notions qui ne semblent évidentes que quand on n'y pense pas. A l'écrit, aucun problème : est un mot toute suite de lettres précédée et suivie par un blanc typographique. Mais nous nous intéressons à l'oral. Là, on entend souvent que le mot est un objet difficile à définir, car les mots s'y enchaînent les uns aux autres, et le FP est sans doute l'une des langues au monde où ces chaînes phonologiques sont les plus serrées : /ʃɛpa/, est-ce un mot ou bien trois comme sa traduction graphique <je sais pas> ? La même incertitude peut du reste affecter l'écrit : <moulin à café> compte trois mots selon le critère des blancs ; mais ce qui est ainsi désigné ne s'en conçoit pas moins comme un simple objet au même titre que la casserole, qui se satisfait d'un mot unique. Pourquoi /mulëakafe/ n'en serait-il pas un à l'instar de son équivalent finnois *kahvimylly*, unique pour la graphie comme pour la prononciation ?<sup>21</sup>

A ces interrogations, on peut répondre de plusieurs façons. Premièrement, il est bien clair que les enfants acquérant le langage acquièrent aussi, voire surtout, des mots, de un ou deux à 18 mois jusqu'à plusieurs milliers dès l'âge de six ans (voir Kail & Fayol 2000). Et même s'il est vrai que les textes antiques ne séparent pas toujours les mots, l'invention très ancienne de moyens graphiques de le faire, point médian puis espace, suffit à montrer que le concept n'est pas artificiel, mais repose bien sur une propriété naturelle du langage humain. La notion de mot composé — <moulin à café> en est un — permet de dépasser l'alternative du un et du plusieurs.<sup>22</sup>

Deuxièmement, il existe des indices neurologiques sérieux de la réalité de la notion de mot. L'expérience que je vais décrire met en jeu des mots écrits, en sorte qu'elle n'est pas tout

<sup>21</sup> La force et l'arbitraire de la connexion des mots aux choses sont ici déterminants. Enfant, j'ai longtemps entendu /vaʃkibi/ comme le nom indécomposable d'un fromage en portions, sans faire le moindre rapprochement avec les bovins et l'hilarité.

<sup>22</sup> L'orthographe n'est guère cohérente qui tantôt sépare les parties du composé par des blancs (<moulin à café>), tantôt les lie par des tirets (<tire-bouchon>).



à fait concluante, mais elle n'en est pas moins révélatrice d'une inscription neuronale. On présente des mots écrits à une patiente (anglophone) atteinte d'une lésion de l'hémisphère gauche provoquant une négligence spatiale unilatérale de la partie droite du champ visuel, si bien qu'elle a tendance à lire mal ou pas du tout la fin des mots : pour *house* 'maison', elle lit *hound* 'chien de chasse', pour *stripe* 'rayure', elle lit *strip* 'bande', etc. On lui présente ensuite les mêmes mots en miroir (*esuoh*, *epirts*) — pour constater qu'elle éprouve toujours les mêmes difficultés à en lire la fin, bien que celle-ci se trouve à présent à gauche ; disposer les mots verticalement n'a pas davantage d'effet. Comme l'écrit l'auteur qui rapporte cette expérience : « Ce qu'implique ce résultat... c'est que... les mots écrits... doivent être représentés ... quelque part dans son cerveau [de la patiente] comme s'ils étaient écrits de gauche à droite » (Glynn 1999 : 410).<sup>23</sup>

Une étude neurolinguistique récente du traitement phonétique (Gwilliams et al. 2022) va, me semble-t-il, dans le même sens. Je cite : « Nos analyses montrent que le cerveau encode continûment en parallèle les trois derniers sons du langage entendus, et qu'il conserve cette information longtemps après qu'elle a disparu de l'input sensoriel. Chaque représentation de sons du langage évolue avec le temps, encodant à la fois les traits phonétiques et la durée écoulée depuis le début. Le réseau neuronal dynamique ainsi établi encode à la fois l'ordre relatif et le contenu phonétique de la séquence parlée. Ces représentations s'activent d'autant plus tôt que les phonèmes sont plus prédictibles, et elles se maintiennent plus longtemps quand l'identité lexicale est incertaine. »<sup>24</sup> Autrement dit, on a affaire à un traitement « glissant » tel que l'accès à la représentation lexicale s'ouvre dès que le complément de la suite phonétique déjà traitée devient pleinement prédictible. Voilà qui veut bien dire que ladite représentation lexicale a une existence neuronale.

Enfin, quand on reproche aux mots leur difficile délimitation à l'oral, c'est qu'on introduit dans la discussion une nouvelle notion : le mot phonologique, c.-à-d. une séquence sonore qui, en énonciation naturelle, doit se prononcer sans pause interne. Or, il n'existe a priori aucune correspondance entre le mot phonologique et une quelconque unité d'analyse morphosyntaxique : le fameux *Doukipudonktan* ? 'D'où qu'il pue donc tant ?' de Queneau est un mot phonologique et une phrase interrogative. Par là, c'est un *signe* au sens saussurien, c'est-à-dire l'association d'une forme (complexe) et d'un sens (complexe). Mais, du fait même de leur complexité, l'analyse morphosyntaxique de ces séquences-signes — fréquentes en FP pour des raisons de prosodie — est toujours possible, jusqu'à un niveau où il devient possible d'associer des formes simples, non décomposables plus avant, à des significations simples — relativement à la finesse atteignable par nos analyses sémantiques ou souhaitable eu égard à une tâche particulière. Dans la phrase de Queneau, la décomposition mène à six ou sept formes simples : six si l'on s'arrête à *du* 'd'où', *k(ə)* 'qu(e)', *i(l)* 'il', *py* 'pue', *dɔ̃k* 'donc' et *tɑ̃* 'tant' ; sept si l'on considère que *du* combine deux éléments, la préposition *d(ə)* 'd(e)' et l'adverbe *u* 'où'. Décomposer *py*, en revanche, ne mène qu'à deux phonèmes /p/ et /y/ qui ne veulent rien dire, mais sont comme des pièces de lego par rapport aux maisons, avions, etc. qu'elles permettent de construire, ou comme les atomes d'hydrogène et d'oxygène par rapport

<sup>23</sup> « *The implication of this finding is that... the printed words... must be represented... somewhere in her brain as if they were written from left to right* ». L'auteur précise que cette patiente est assez exceptionnelle. La plupart des personnes souffrant du même handicap ne « redressent » pas les mots : quand on les leur présente à l'envers, elles ont du mal avec le début. Ce qui suggère, sans guère nous avancer, que la lésion neurologique peut affecter le traitement lexical à plusieurs niveaux.

<sup>24</sup> « *Our analyses show that the brain continuously encodes the three most recently heard speech sounds in parallel, and maintains this information long past its dissipation from the sensory input. Each speech sound representation evolves over time, jointly encoding both its phonetic features and the amount of time elapsed since onset. As a result, this dynamic neural pattern encodes both the relative order and phonetic content of the speech sequence. These representations are active earlier when phonemes are more predictable, and are sustained longer when lexical identity is uncertain.* » Les durées se mesurent en millisecondes.

à la molécule d'eau H<sup>2</sup>O.<sup>25</sup> En cela, *py* /py/ <pue> est un mot. (Autre affaire est le fait que, pour des raisons sur lesquelles je reviendrai, il ne s'emploie jamais seul mais associé, par exemple, à un pronom sujet (*i(l)=py* 'il pue'). Font exception les propositions « métalinguistiques » comme celle de la phrase précédente ou bien les énoncés impératifs tels que *py !* 'pue !')

Par conséquent, et moyennant la réserve exprimée à la note 5, la distinction la plus basique au sein de l'ensemble « mots » est sans doute celle entre mots simples et mots complexes. *Kasəɾɔl* 'casserole' est un mot simple : on n'y peut isoler aucune sous-partie signifiante, on ne peut que le casser en six ou sept phonèmes, selon que l'on prononce ou non le schwa — jamais dans mon parler francilien, variablement en terre occitane. *Mulɛ̃#a#kafɛ* 'moulin à café' est un mot composé, donc complexe. Mais, de façon moins évidente quoique indiscutable, *aspir#atəɾ* 'aspirateur' l'est aussi, qui désigne un instrument comme beaucoup de noms terminés en /(at)œʁ/ — cf. *miks#əɾ* 'mixeur', *ordin#atəɾ* 'ordinateur', etc. —<sup>26</sup> et en même temps indique explicitement (ce que ne fait pas *kasəɾɔl* 'casserole') à quoi sert l'instrument, à savoir aspirer (la poussière entre autres), si bien qu'on peut en effet y déceler deux parties. Et de même, peut-être encore moins évident mais tout aussi sûr, *mul-ɛ* 'moulaie | moulaie | moulaie', puisque, d'une part, le mot fait référence à une certaine activité, et d'autre part, il donne à entendre que ladite activité relève d'une période antérieure au moment où il est articulé (imparfait) et a été effectuée ou bien par qui parle (1<sup>ère</sup> personne du singulier), ou bien par à qui on parle (2<sup>e</sup> personne du singulier), ou bien par une ou plusieurs tierces personnes (3<sup>e</sup> personne du singulier ou du pluriel). On verra plus loin et au chapitre 9 comment ces ambiguïtés se résolvent et comment les deux types de significations (nature de l'activité, temps et personne) s'associent aux deux parties de la forme : /mul/ et /ɛ/.

Il est bien rare qu'une distinction binaire (une dichotomie) ne laisse rien échapper. Le reste est ici fait de mots comme *mulɛ̃* 'moulin' et *banal* 'banal'. Le premier est sémantiquement complexe puisqu'il désigne, à la façon d'*aspir#atəɾ* 'aspirateur', une machine servant à moulin. Du point de vue morphologique, en revanche, il est simple. La finale /ɛ̃/ ne peut être analysée comme une partie régulièrement associée à la signification « instrument » à la manière de /(at)œʁ/. (On verra ci-dessous ce qu'il faut entendre par « régulièrement ».) Quant à *banal*, les historiens de la langue savent que cet adjectif est dérivé de *ban*, terme médiéval désignant l'ensemble des vassaux d'un seigneur. Mais ils sont les seuls. En français contemporain, le mot n'apparaît plus guère que dans l'expression « publier les bans », avec laquelle *banal* a coupé tout lien sémantique, en sorte qu'il s'agit bien à présent d'un mot simple — sans que ce soit du reste le cas de tous les adjectifs en /al/ : cf. *āvirɔnəmātal* 'environnemental' en regard de *āvirɔnəmā* 'environnement'. On voit qu'analyse morphologique et histoire de la langue (diachronie) ne se laissent pas séparer.

Cette ébauche de classement nous amène à descendre d'un échelon dans la hiérarchie et à poser une distinction, moins radicale que celle entre simples et complexes, au sein des mots complexes. Laissant de côté les mots composés (*mulɛ̃#a#kafɛ* 'moulin à café', *tir#buʃɔ̃* 'tire-bouchon'), je veux parler de la différence entre *aspir#atəɾ* 'aspirateur' et *mul-ɛ* 'moulaie | moulaie | moulaie'. A première vue, ces deux types de mots complexes se ressemblent, formés qu'ils sont de deux parties dont l'une renvoie à une activité déterminée, aspirer et moulin. Sauf qu'*aspir#atəɾ* 'aspirateur', à la différence de *mul-ɛ* 'moulaie | moulaie | moulaie |

<sup>25</sup> A moins que *py* ne « contienne » des éléments inaudibles mais dotés de significations, selon une hypothèse que je discuterai plus loin. Que les mots signifiants sont constitués d'un ensemble fini d'éléments non signifiants (les phonèmes) se nomme la double articulation, et c'est une propriété fondamentale du langage humain, que les signaux communicatifs d'autres espèces animales (chimpanzés, dauphins, etc.) ne présentent pas, semble-t-il. Il peut évidemment se trouver qu'un seul phonème suffise à former un mot : cf. /o/ o <eau>.

<sup>26</sup> Beaucoup, mais évidemment pas tous : cf. *fāt#əɾ* 'chanteur' et *ʃaləɾ* 'chaleur'. Le premier, complexe à la façon d'*aspir#atəɾ* 'aspirateur', dénote une personne (ou un oiseau) qui chante, un agent en termes techniques ; le second est inanalysable.

moulaient’, ne dénote pas l’activité, mais l’instrument qui permet de s’y livrer. Par rapport à *aspir-e* ‘aspirer’ *aspir#atær* ‘aspirateur’ accomplit donc une transmutation ontologique — un changement dans la nature des choses, pardon pour le jargon — de procès à objet, qui se traduit grammaticalement par le fait qu’*aspir-e* ‘aspirer’ est un verbe, *aspir#atær* ‘aspirateur’ un nom, celui-ci « dérivé » de celui-là. Rien de tel avec *mul-ε* ‘moulaie | moulaie | moulaie’ : il s’agit bien de moudre, mais spécifié pour qui le fait et quand. Ce qui vient d’être là caractérisé, c’est la distinction entre deux types de morphologie : dérivationnelle (ou constructionnelle), qui forme des mots à partir d’autres en en changeant plus ou moins profondément la référence ; flexionnelle, qui modifie la référence selon certaines dimensions, sans la changer ; et entre deux types de mots complexes : les mots dérivés et les mots fléchis.<sup>27</sup> Seuls les seconds nous occuperont désormais. (Pour les premiers, voir Fradin 2003 ; Haspelmath & Sims 2010.)

Qu’est-ce exactement qu’un mot fléchi ? A un certain niveau d’abstraction, c’est-à-dire d’imprécision acceptable, il est permis d’employer le terme « mot » comme je l’ai fait jusqu’ici pour désigner conjointement les *lexèmes* et les *mots-formes* qui les manifestent. Il est temps de débrouiller la confusion. Le lexème est une entité abstraite, la dénomination attribuée à l’ensemble des mots-formes correspondant à une entrée lexicale unique, lesquels ont seuls une réalité concrète (sonore ou graphique). Par exemple, le lexème latin PONS ‘pont’ — conventionnellement noté en petites capitales et en orthographe comme expliqué au chapitre précédent — se manifeste par dix mots-formes réalisant cinq formes casuelles (nominatif, accusatif, génitif, datif, ablatif) aux deux nombres (singulier et pluriel), p.ex. *pons* (nominatif singulier), *pontis* (génitif singulier), *pontibus* (datif-ablatif pluriel), etc. Le lexème français VACHE se réalise en FP par le seul mot-forme *vaf* /vaf/, mais par les deux mots-formes graphiques <vache> et <vaches> en FO. De même encore, les mots-formes *dør*, *dørmir*, *dørmi-ǝ*, etc. sont autant de réalisations du lexème DORMIR, le choix de l’infinitif pour le représenter purement conventionnel, comme on sait.<sup>28</sup> (Pour la notion de « mot-forme », cf. Fradin 2003 : 85 ; Haspelmath & Sims 2010 : 15-16.) L’ensemble des mots-formes réalisant un lexème constitue le *paradigme* du lexème.

On notera au passage qu’il n’est pas requis que tous les mots-formes d’un paradigme soient matériellement distincts les uns des autres. Ainsi, au pluriel, les formes nominative (cas du sujet) et accusative (cas de l’objet direct) de PONS sont homophones et homographes, *pontes*. Une décision s’impose alors. Ou bien on pose l’existence de deux mots-formes exprimant deux propriétés morphosyntaxiques distinctes (ou davantage) par le biais de formes sonores et graphiques accidentellement identiques ; ou bien d’un seul mot-forme *synchrétique*, c’est-à-dire exprimant indistinctement les deux propriétés (ou davantage).<sup>29</sup> Le choix ne saurait être arbitraire, mais devra s’appuyer sur les significations respectives des propriétés et la probabilité qu’elles puissent se confondre, et il ne manquera pas non plus de tenir compte de l’histoire des formes. Ainsi, notre exemple latin représente probablement un cas de *synchrétisme*, pour autant que nominatif et accusatif partagent la propriété de marquer les arguments les plus directs du verbe, le sujet et l’objet direct, autrement dit les participants quasi obligés du procès dénoté par un verbe transitif. On ne s’étonnera donc pas d’une forme cumulant l’expression des deux fonctions grammaticales.<sup>30</sup> Au contraire, l’identité formelle de *(ty=)ǝtæra* ‘tu chanteras’ et *(el=)ǝtæra* ‘(elle) chantera’ en FP a toutes chances d’être

<sup>27</sup> Bien évidemment, un mot dérivé peut être fléchi, p.ex. *banal#izæ-rǝ* ‘banaliserons’, forme fléchie (1<sup>ère</sup> personne du pluriel du futur) du verbe *banal#iz-e* ‘banaliser’, lui-même dérivé de l’adjectif *banal* ‘banal’.

<sup>28</sup> Le latin préfère la 1<sup>ère</sup> personne du singulier du présent de l’indicatif, d’où DORMIO.

<sup>29</sup> « Synchrétisme... est employé pour désigner le phénomène par lequel une forme se trouve appelée à cumuler plusieurs fonctions » (Marouzeau 1961 : 220).

<sup>30</sup> Mais on ne s’étonnera pas davantage, même si on n’en comprend pas la raison, que ce cumul ne s’observe qu’au pluriel : cf. nominatif singulier *pons* vs. accusatif singulier *pontem*. La cohérence n’est pas une caractéristique des langues naturelles.

accidentelle. On discerne mal quelle cause sémantique amènerait la 2<sup>e</sup> et la 3<sup>e</sup> personne du singulier à se confondre — abstraction faite du pronom personnel, dont la présence est évidemment cruciale (voir le chapitre 9).<sup>31</sup> On sait en outre que, jusqu'à une date relativement récente (seizième siècle), les deux formes étaient distinctes, car le /s/ final de la 2<sup>e</sup> personne se prononçait encore, avant de s'amuir en compagnie de toutes les consonnes finales (cf. Gougenheim 1951).<sup>32</sup>

Un paradigme dont chaque case est distincte de toutes les autres est dit « canonique » (cf. Corbett 2009). Ainsi le paradigme du présent de l'indicatif des verbes du 1<sup>er</sup> groupe en espagnol, soit pour *cantar* 'chanter' : *canto*, *cantas*, *canta*, *cantamos*, *cantais*, *cantan* 'je chante, etc.' Plutôt que la règle, la canonicité constitue l'exception. Cela se comprend, s'il est vrai qu'un canon est une somme de « règles déterminant un idéal » (Rey et al. 2013 : 101).

### 5.3. Mots pleins, mots fonctionnels, affixes

On veut parler ici d'une autre distinction que celle entre mots simples et mots complexes, ou plutôt de deux distinctions qui se recoupent en partie : d'une part entre mots pleins (*content words*) et mots grammaticaux ou fonctionnels (*function words* ; cf. Creissels 2006a : 18-19) ; d'autre part, entre mots, pleins ou fonctionnels, et affixes. Les deux distinctions, simple vs. complexe et mots (pleins vs. fonctionnels) vs. affixes, sont mutuellement indépendantes pour autant que mots pleins et mots fonctionnels peuvent aussi bien être simples que complexes : *mulē* 'moulin' est plein et simple, *mul-jō* 'moulin' plein et complexe, *kōtr* 'contre' fonctionnel et simple, *le* 'les' fonctionnel et complexe (car cumulant les propriétés « défini » et « pluriel »). Les affixes sont eux aussi l'un ou l'autre : le /-s/ de l'occitan languedocien *cabras* 'chèvres' est simple, car il n'encode que la pluralité ; le /-jō/ de *mul-jō*, si on le compte pour un, est complexe, exprimant à la fois le temps du verbe (imparfait) et la personne du sujet (1<sup>ère</sup> du pluriel).

« Plein », pour être courant, n'est guère satisfaisant, de même d'ailleurs que l'anglais *content*, car les mots fonctionnels ne sont sûrement pas vides de contenu. Je lui préférerais désormais « référentiel », considérant que les mots de cette sorte ont pour caractéristique de faire référence à des phénomènes du monde réel — ce à quoi « on se cogne », comme disait Jacques Lacan — ou bien mental et imaginaire, sans qu'il y ait évidemment de cloison étanche entre l'un et l'autre.<sup>33</sup> Les mots fonctionnels spécifient de diverses manières la référence des mots référentiels : par exemple *la=lāp* 'la lampe' où l'article défini *la* exprime les propriétés [DÉFINITUDE +], [NOMBRE *singulier*] et [GENRE *féminin*] (voir le chapitre 6) ; *ty=py* 'ty pues', où *ty* 'tu' « désigne » quiconque s'entend ainsi apostrophé et n'est donc associé à aucun référent stable. On est donc tenté de considérer qu'ils ne réfèrent pas. L'affirmer serait toutefois un peu trop péremptoire. Certaines prépositions, par exemple, ont un sens locatif ou temporel assez concret, qui semble bien représenter une espèce de référence : cf. *ā* 'en' et *pādā* 'pendant' dans *il=a\_kury lə=maratō ā=dōz=ōr* 'Il a couru le marathon en deux heures', qui dit bien qu'il a fait ses quarante kilomètres, et *il=a\_kury lə=maratō pādā=dōz=ōr* 'Il a couru le marathon pendant deux heures', qui suggère plutôt qu'il s'est arrêté avant la fin. En outre, il est connu que d'assez nombreuses langues amérindiennes d'Amérique du Nord présentent des éléments qu'on ne peut que qualifier de fonctionnels, mais qui ont « des significations habituellement exprimées par des mots séparés dans les langues européennes qui nous sont plus familières » (Mithun 1999 : 48).<sup>34</sup> Il y a ainsi

<sup>31</sup> Stump (2001) distingue le syncrétisme motivé, type *pontes*, du syncrétisme arbitraire, type *fātara*. Il me paraît préférable, suivant la définition de Marouzeau, de réserver le terme à la première sorte.

<sup>32</sup> La liaison de /tyfātərazavəkɪmwa/ pour *ty=fātə-ra avək=mwa* 'tu chanteras avec moi', qui semble faire ressurgir le /s/ amui, ne s'entend pas en FP courant ; elle est probablement due à l'influence de l'orthographe.

<sup>33</sup> Et compte tenu des homonymies, sens dérivés, métaphores, etc. Ce renvoi à un *réfèrent* (cf. Ogden & Richards 1923 ; Dubois et al. 1973 : 414-415) est bien entendu médiatisé par un concept, le *signifié* saussurien.

<sup>34</sup> « *meanings that are usually expressed in separate words in the more familiar languages of Europe* ».

en nisgha, langue tsimshian de Colombie Britannique, parmi des dizaines d'autres, un préfixe ou proclitique (voir ci-dessous) *q'altix* qui, attaché à un verbe, ajoute le sens de « dans/vers la forêt », p.ex. *q'altix=yé* 'aller dans la forêt', sans pourtant exister en tant que mot référentiel signifiant « forêt » (Mithun 1999 : 145).<sup>35</sup> On notera quand même, d'une part, que ce phénomène semble rare dans les langues du monde ; d'autre part, que la différence entre référence autonome et signification adventice ou secondaire, éventuellement quasi référentielle, ne semble pas être mise à mal par ces cas particuliers.

Ou bien les mots fonctionnels tiennent la place de mots référentiels, auquel cas leur interprétation dépend entièrement du contexte discursif et/ou physique : p.ex. *səlyisi* 'Celui-ci' en réponse au *vu=preferē\_kel=məlō* ? 'Vous préférez quel melon ?' du commerçant. Cela étant, pas davantage que le réel et l'imaginaire, le référentiel et le fonctionnel ne sont séparés par une frontière infranchissable. Il est assez banal de voir un mot référentiel utilisé sur le mode fonctionnel. Le verbe AVOIR est référentiel quand il désigne une certaine relation entre des entités : p.ex. *ma=vwazin a ē=bo\_fa* 'Ma voisine **a** un beau chat' ; il est fonctionnel dans la périphrase du passé composé, p.ex. *ma=vwazin a\_perdy sō=fa* 'Ma voisine **a** perdu son chat'. (Quant aux périphrases, voir plus loin.) En wolof, *biir* /bi:r/ veut dire 'ventre', mais la locution prépositionnelle *ci biir* /cibi:r/, littéralement 'dans ventre', ne signifie rien de plus que 'dans', à moins que le contexte n'en décide autrement : comparez *ci biiru kër gi* 'dans la maison' et *ci biiru nag bi* 'dans le ventre de la vache' — quoique 'dans la vache' ne puisse être exclu.

Non contente d'être assez évidente au vu des données linguistiques, la différence entre mots référentiels et mots fonctionnels est neurologiquement fondée. Le démontrent les patients atteints d'une aphasie de Broca, dont le discours se caractérise par la quasi absence des seconds (articles, pronoms, etc.) et, quant aux premiers, une prépondérance de noms (fortement référentiels), moins de verbes, rarement fléchis, et peu d'adjectifs. L'aphasie dite de Wernicke montre la même chose à l'inverse : les mots fonctionnels sont bien préservés, le déficit touche l'emploi des mots référentiels (cf. Sabouraud 1995).<sup>36</sup>

Dans la mesure où les mots fonctionnels se fléchissent, ils relèvent comme les mots référentiels de la distinction entre lexème et mot-forme. On peut ainsi considérer que l'article défini du français consiste en un lexème L, dont les mots-formes sont *lə* 'le', *la* 'la' et *le* 'les' (cf. chapitre 6). En grec moderne, le même article défini, fléchi pour trois genre (masculin, féminin, neutre), deux nombres (singulier et pluriel) et trois cas (nominatif, accusatif, génitif), se réalise par 18 mots-formes, dont seuls dix sont phonologiquement distincts (Mirambel 1939/2002 : 40). De ces 18 formes, 14 commencent par /t/, si bien que T peut être une notation adéquate du lexème. En anglais, la distinction n'est que de notation : THE pour le lexème, /ðə/ (/ðr/ devant voyelle) pour l'unique mot-forme le réalisant.

Les mots référentiels se répartissent entre plusieurs parties du discours (aussi nommées « catégories lexicales ») selon des prototypes sémantiques (cf. Croft 1991 ; Creissels 2006a ; Abeillé & Godard 2021). Les noms communs dénotent prototypiquement des entités limitées dans l'espace et stables dans le temps (des choses), p.ex. *lāp* 'lampe'.<sup>37</sup> Mais un prototype est comme un canon un idéal voué aux écarts, de fait nombreux, p.ex. les abstraits comme *liberte* 'liberté', les noms d'action comme *destryksjō* 'destruction', etc. Les verbes dénotent prototypiquement des entités non limitées dans l'espace et changeantes dans le temps, c.-à-d. des états mentaux ou physiques (*savwar* 'savoir', *gralote* 'grelotter'), des procès (*myrir* 'mûrir'), ou bien des événements et des actions (*neze* 'neiger', *kurir* 'courir'). Les adjectifs dénotent prototypiquement des propriétés attribuées aux entités dénotées par des noms

<sup>35</sup> Autrement dit, *q'altix=yé* n'est pas un mot composé à la façon de l'anglais *kickstart* 'démarrer au pied' (une moto).

<sup>36</sup> Que se passe-t-il quand une locutrice ou un locuteur du nisgha est atteint.e d'aphasie de Broca ou de Wernicke ? On aimerait le savoir.

<sup>37</sup> J'adopte ici l'ontologie de Zemach (1970, 1979).

(*lāp\_zon* ‘lampe jaune’). Les adverbes font de même vis-à-vis de ce que dénotent les verbes (*kurir rapidamā / vit* ‘courir rapidement / vite’) ou des situations décrites par des prédicats (*I(l)=než dāṛ* ‘Il neige dehors’).

Ce tableau est évidemment très insuffisant, d’autant que la classification elle-même présente des lacunes : par exemple, les numéraux sont-ils des adjectifs cardinaux (*le=trwa\_muskāter* ‘les trois mousquetaires’ comme *le=vjō\_muskāter* ‘les vieux mousquetaires’) ? Ou bien des noms (*Done=mwa la=trwa* ‘Donnez-moi la trois’ comme *Done=mwa la=fābrā\_trwa* ‘Donnez-moi la chambre trois’) ? Il semble bien qu’il faille en faire une catégorie sémantique à cheval sur plusieurs parties du discours. Il en va du reste ainsi des quantificateurs en général : cf. *boku* ‘beaucoup’ dont le comportement est tantôt celui d’un nom (*boku diz kə...* ‘beaucoup disent que...’ comme *le=zā diz kə...* ‘les gens disent que...’), tantôt celui d’un adverbe (*I(l)=māz boku* ‘Il mange beaucoup’ comme *I(l)=māz vit* ‘Il mange vite’).

Les mots fonctionnels appartiennent eux aussi à diverses catégories, plus ou moins bien définies. Le FP connaît les suivantes : conjonctions (*ε* ‘et’, *mε* ‘mais’, etc.), déterminants (*la* ‘la’, *sə* ‘ce’, *mō* ‘mon’, etc.), prépositions (*dā* ‘dans’, *pur* ‘pour’, etc.), pronoms (*zə* ‘je’, *mwa* ‘moi’, *ki* ‘qui’, etc.), subordonnants (*kā* ‘quand’, *kə* ‘que’, *si* ‘si’, etc.).

Non contentes de reposer sur des fondements sémantiques ou cognitifs différents, les parties du discours majeures (noms, verbes, adjectifs) se distinguent aussi, et plus fermement, par le fait de relever le plus souvent de types flexionnels bien distincts. Le latin est exemplaire à cet égard, où les flexions nominales et verbales sont à peu près entièrement disjointes.<sup>38</sup>

Du point de vue morphologique, les mots référentiels sont libres en ce sens qu’ils peuvent en principe s’énoncer seuls — rarement en FP — et que leur position par rapport à d’autres constituants n’est pas fixe, mais dépend entre autres choses de leur fonction syntaxique (sujet, objet, etc.). Les mots fonctionnels sont libres (p.ex. *selsi* ‘celle-ci’) ou liés (*bound*) à tels constituants. Si liés, on les appelle des « clitiques », ainsi définis par Abeillé & Godard (2021, Glossaire) : « Forme généralement monosyllabique, qui doit se combiner avec un élément appelé *hôte* pour former un GROUPE 1 prosodique, par exemple *je* ou *-ci* ». <sup>39</sup> Les autrices retiennent ainsi avant tout le critère phonologique, à savoir que les clitiques sont inaccentués et doivent s’appuyer sur un mot accentué, ce pourquoi ils ne s’énoncent jamais seuls, sauf en discours métalinguistique (« *je* est un pronom »). Ce critère, issu de la grammaire des langues classiques, du grec en particulier, s’applique bien aux langues romanes autres que le FP, mais mal à celui-ci. Car, on le sait, l’accent tonique n’est pas en FP une propriété des mots, mais de suites de mots qui peuvent coïncider — mais pas toujours — avec des syntagmes, dont il frappe la syllabe finale. Il s’ensuit que le clitique se trouve accentué dans *U\_va=ty ?* ‘Où vas-tu ?’ et *Di=lə !* ‘Dis-le !’, prononcés /uva’ty/ et /di’lə/.<sup>40</sup>

Il existe deux sortes de clitiques, qu’il convient de bien distinguer, à savoir les clitiques simples et les clitiques spéciaux.<sup>41</sup> Les premiers sont en fait des mots pleins, mais phonétiquement réduits au point de perdre leur autonomie prosodique. L’anglais en offre des exemples, p.ex. *I’m* pour *I am* ‘je suis’, où *m* est la forme clitique de *am* ; le latin parlé tel que nous le restitue Plaute également : cf. *Illuc homo superstitiosu-st* {ce homme sorcier-est} ‘Cet

<sup>38</sup> Le participe passif (p.ex. *amatus* ‘aimé’) fait exception, qui est une forme verbale, mais se fléchit comme un adjectif. Mais les participes sont notoirement des formes mixtes, verbo-nominales.

<sup>39</sup> Un groupe 1 prosodique est un groupe de mots ne comportant qu’un seul accent d’intensité principal. Le mot a été forgé par le linguiste américain Eugene A. Nida (1914-2011) à partir du grec ancien *κλινῶ* voulant dire entre autres choses « s’appuyer contre » (cf. Nida 1949).

<sup>40</sup> En fait, *lə* ‘le’ fait exception. L’autre pronom en /ə/ susceptible d’être enclitique, *zə* ‘je’, perd le schwa dans cette position : cf. *u\_syi=zə ?* /usɥiz/ ‘Où suis-je ?’ (Milner & Regnault 1987 : 86-87). Quant à *sə* ‘ce’, il n’est tonique que dans la locution *syr=sə* /syksœ/ ‘sur ce’. L’ouvrage de référence sur la phonétique française reste Delattre (1966).

<sup>41</sup> Je reprends la terminologie de Zwicky (1977). Voir aussi Creissels (2006a : 30-33).

homme est sorcier’, *-st* la forme clitique de *est* ‘est’ (*Amphitryon*, Morisset & Thévenot 1975 : 33). Le français les ignore, semble-t-il. Les clitiques spéciaux, auxquels ne correspond aucune forme non-clitique de contenu identique, sont, quant à eux, des mots fonctionnels, porteurs de ces significations secondaires qui les séparent des mots référentiels. C’est d’eux seulement qu’il sera question dans la suite de l’étude, si bien qu’il ne sera plus nécessaire d’en préciser l’identité.

Selon leur position devant ou derrière l’hôte, les clitiques sont dits proclitiques, par exemple les pronoms objets de *marsəlin lə=lyi=di* ‘Marceline le lui dit’ ; ou enclitique comme ce même pronom avec un verbe à l’impératif : *di=lə=lyi* ! ‘Dis-le-lui !’. On voit que la propriété positionnelle n’est pas nécessairement inhérente, mais peut dépendre du contexte. Le fait est néanmoins qu’elle l’est le plus souvent : l’article défini est proclitique en français, enclitique en roumain (*calul* /kal=ul/ ‘le cheval’). Il est aussi d’usage de distinguer les clitiques de mot (*word clitics*) des clitiques syntagmatiques ou de constituant (*phrase clitics*). L’article défini français, qui se lie, on l’a vu, au premier mot du SN quel qu’il soit, constitue un bon exemple de ces derniers. Un clitique de mot, en revanche, est censé ne se lier qu’au mot qu’il modifie, dont il ne peut être séparé. Il n’est pas sûr que cette distinction soit justifiée (voir ci-dessous).

Mots pleins et mots fonctionnels s’opposent aux affixes en cela que ceux-ci ne sont pas des mots mais, du moins dans le modèle ici appliqué (voir plus loin), des objets phonologiques indissociables des mots qui les contiennent : par exemple /-em/ dans latin *pontem*, accusatif de *pons* ‘pont’, ou encore /-ε/ et /-war/ dans *I(l)=pløv-ε* ‘Il pleuvait’ et *pløv-war* ‘pleuvoir’, imparfait et infinitif du lexème verbal PLEUVOIR. En d’autres termes, on n’attribue pas d’entrée lexicale à /-em/, /-ε/ et /-war/. Il en est une, en revanche, pour les lexèmes PONS et PLEUVOIR et les divers mots-formes qui les réalisent (leurs paradigmes), ainsi que pour les mots fonctionnels, libres ou liés, p.ex. LUI ‘3SG.M’, dont le paradigme consiste en quatre mots-formes, dont trois clitiques : *i(l)*= ‘3SG.M.SUJ’, *lə*= ‘3SG.M.OD’, *lyi*= ~ *i*= ‘3SG.M.OI’ et *lyi* ‘3SG.M’ (cf. chapitre 6).

Affixes et clitiques (spéciaux) partageant les deux propriétés d’exprimer des significations fonctionnelles — le plus souvent, à l’exception des langues amériindiennes mentionnées plus haut — et d’être liés, les distinguer ne va pas toujours sans quelque flou. Trois critères se révèlent en général efficaces : (a) les affixes peuvent consister en un unique phonème (p.ex. *-t* dans *vu=di-t* ‘vous dites’), voire en une modulation accentuelle ou tonale, comme on le verra à la section suivante, alors que les clitiques, quoique brefs, ont d’ordinaire davantage de substance phonétique (ce sont souvent d’anciens mots libres) ; (b) la présence d’un affixe donné peut être corrélée à une forme particulière de la base — la flexion verbale du FP en offre, on le verra, de nombreux exemples — alors que les clitiques s’attachent à la base à la forme qu’elle présente en leur absence ;<sup>42</sup> (c) les affixes ne se lient qu’à une catégorie lexicale déterminée, propriété que ne partagent pas les clitiques ; (d) les affixes ne sont pas mobiles par rapport à la base, les clitiques peuvent l’être.<sup>43</sup>

Ces deux dernières propriétés sont sans doute les plus discriminantes. A première vue, l’article défini roumain illustré plus haut (*calul* ‘le cheval’ vs. *un cal* ‘un cheval’) ressemble à un suffixe — c’est d’ailleurs ainsi qu’il est souvent décrit. En fait, de même que pour son équivalent français, l’hôte n’est pas le nom auquel est attribué la propriété d’être défini, mais le premier mot du SN dont ce nom est la tête : cf. *frumosul cal* ‘le beau cheval’ vs. *calul alb* ‘le cheval blanc’. Etant donné (c), on a donc bien affaire à un clitique spécial syntagmatique. La condition (d), pour sa part, suffit à montrer que les pronoms objets « atones » du français,

<sup>42</sup> Moyennant quelque variation. Ainsi, le clitique latin =*que* ‘et’ attire l’accent sur la syllabe qui le précède : cf. *populusque* [popuˈluskwe] ‘et le peuple’ vs. *populus* [ˈpopulus] ‘peuple’.

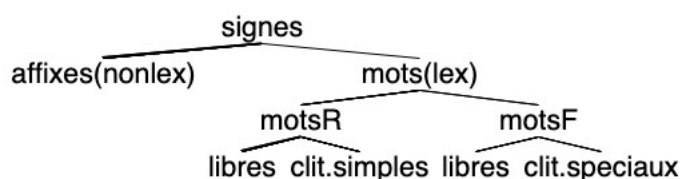
<sup>43</sup> « *the base of a morphologically complex word is the element to which a morphological operation applies* » (Haspelmath & Sims 2010 : 321). Ce terme, pratique parce que général, sera précisé plus loin.

uniques ou bien formant un groupe comptant pour un seul élément (voir les exemples ci-dessus), sont bien des clitiques et non des affixes, en dépit du fait qu'ils sont généralement adjacents au verbe, mais pas toujours du même côté

Il est des langues, en revanche, où la marque de définitude semble clairement affixale. En arabe classique, /l-/ s'attache au nom défini ainsi qu'aux éventuels adjectifs épithètes qui s'accordent en définitude comme ils le font pour les autres traits flexionnels : cf. *l-baab-u l-kabiir-u* {DEF-porte.M-NOM DEF-grand.M-NOM} 'la grande porte'.<sup>44</sup> Y voir un préfixe paraît donc s'imposer. En islandais familier, dans un SN adjectif-nom, le déterminant défini se suffixe à la tête nominale, l'accord en définitude seulement marqué par la déclinaison dite faible de l'adjectif: cf. *unga konan* /'uŋ-a 'k<sup>h</sup>o:na-n/ {jeune-NOM.FAIBLE.F femme-DEF.NOM.F} 'la jeune femme', à comparer avec *ung kona* 'une jeune femme' où l'adjectif présente la déclinaison dite forte, seule marque apparente de l'indéfinitude du SN (Pétursson & Eyjólfsson 1996 : 55).<sup>45</sup>

Les pronoms objets clitiques du français sont-ils de mots ou syntagmatiques ? Deux faits plaident en faveur de la seconde hypothèse. Premièrement, avec une forme verbale périphrastique, les pronoms se lient, non au verbe principal dont ils expriment l'objet, mais à l'auxiliaire qui peut, quant à lui, en être séparé : *marsəlin lə=lyi=a pətetr di* 'Marceline le lui a peut-être dit'. Deuxièmement, dans une variété à présent archaïque mais toujours utilisable de la langue, les pronoms objets peuvent « monter » par dessus un verbe modal : cf. *marsəlin lə=lyi=vule dir* 'Marceline le lui voulait dire'. Avec une construction causative, cette « montée » est obligatoire : *marsəlin lə=lyi=fəra repare* 'Marceline le lui fera réparer' — mais c'est qu'il s'agit en fait d'une périphrase (voir section suivante).

L'arbre d'héritage ci-dessous résume ce qui précède (R = référentiel, F = fonctionnel, lex = lexical) :



#### 5.4. Les types de flexion

On entend donc par flexion les divers procédés qui permettent de moduler l'information fournie par un mot sans en altérer la signification fondamentale ni la catégorie lexicale alias partie du discours (nom, verbe, adjectif, etc.). Les modulations ainsi effectuées touchent de nombreuses propriétés, dont certaines, par exemple la quantité, affectent aussi les objets du monde, tandis que d'autres, par exemple la voix active ou passive, sont propres à la cognition humaine et (donc) au langage. Je n'ai ni l'ambition ni la capacité d'en proposer une liste exhaustive, vu leur diversité dans les langues naturelles. En français (FO et FP), ces propriétés sont, pour les noms et les adjectifs, le nombre (singulier vs. pluriel) et le genre grammatical (féminin vs. masculin) ; pour les pronoms personnels, le nombre, le genre grammatical et le cas lié à la fonction grammaticale : cf. *marsəlin lə=ləɾ=a di* 'Marceline le leur a dit', où *lə* /le/ a les propriétés [masculin, singulier, accusatif = objet direct] et *ləɾ* 'leur' les propriétés [féminin/masculin, pluriel, datif = objet indirect] ; pour les pronoms interrogatifs, le caractère humain ou non du référent (*ki* ? 'qui ?' vs. *kwa* ? 'quoi ?' ou *kə* ? 'que ?'). Pour les verbes,

<sup>44</sup> Le caractère défini du SN se marque en outre par le fait que le désinence casuelle (ici -u nominatif) n'est pas suivie du /n/ qu'elle comporterait si le SN était indéfini (*baab-un* 'une porte').

<sup>45</sup> Dans un registre plus littéraire, la présence d'un adjectif entraîne l'emploi de l'article proclitique aux dépens du suffixe : *hin unga kona* {DEF.NOM.F jeune-NOM.FAIBLE.F femme}.



on citera sans plus de commentaire (cf. chapitre 9) le temps, l'aspect, le mode, la voix, la polarité, la personne, le nombre. Ces listes sont assez typiques de ce qu'on observe dans les langues connues, même si la variété n'en reste pas moins importante. Du point de vue des fondements cognitifs de la faculté de langage, il peut être intéressant de remarquer que certaines propriétés évidentes des objets du monde semblent n'être jamais prises en compte par la grammaire (grammaticalisées). Ainsi la couleur : il n'existe pas de langue, semble-t-il, où la flexion marque spécifiquement, disons, les objets rouges par opposition aux objets noirs.<sup>46</sup> Il en est, en revanche, où elle marque la dangerosité des objets (cf. Lakoff 1987).

Depuis les travaux du philosophe et linguiste allemand Friedrich Schlegel (1772-1829), la tradition s'est établie de distinguer deux types de flexion, la synthétique et l'analytique. La flexion synthétique, souvent considérée, à tort, comme la seule vraie flexion, connaît deux procédés. L'un est l'affixation qui, définie de la façon la plus élémentaire possible, consiste à allonger la forme phonologique du mot afin de produire les modulations détaillées plus haut. L'allongement, on l'a vu au paragraphe précédent, se fait au moyen de segments phonologiques, les affixes, plus précisément nommés suffixes ou préfixes selon leur position par rapport à la base.<sup>47</sup> Les exemples simples étant difficiles à trouver en FP (on verra pourquoi), j'en prends un à nouveau chez un proche voisin, l'occitan languedocien : *lauseta* /laʊ'zeto/ 'alouette' et *lausetas* /laʊ'zetos/ 'alouettes' sont deux formes nominales qui désignent (dénotent, réfèrent à) la même entité aviaire, mais on emploie le premier à propos d'un individu unique,<sup>48</sup> le second à propos de plusieurs ; et le procédé se manifeste par la constance de la séquence /laʊ'zeto/, par elle-même associée à la singularité, mais l'apparition à sa suite au pluriel — je m'efforce de m'en tenir à une formulation aussi neutre que possible quant à ce qui se passe « réellement » — du phonème-suffixe /-s/.

En swahili, langue bantoue d'Afrique de l'Est, l'opposition du singulier au pluriel se marque de même en ajoutant à la base. Sauf que l'ajout est un préfixe : cf. *mtoto* / *watoto* 'enfant(s)', /toto/ la partie constante associée au concept « enfant », /m-/ et /wa-/ exprimant simultanément le genre grammatical — ici le genre I réservé aux êtres humains et composé des deux classes nominales numérotées 1(/m-/) et 2 (/wa-/) — et le nombre, singulier pour 1, pluriel pour 2.<sup>49</sup>

Cumuler genre et nombre est assez banal dans les langues pourvues d'une catégorie morphologique de genre. On l'observe, par exemple, en italien : cf. *allodola* / *allodole* 'alouette(s)', où /e/ est la marque de pluriel des noms féminins. Le swahili se distingue apparemment par son double marquage du nombre, alors qu'en occitan et en italien les finales /o/ (<a>) et /a/ ne sont corrélées qu'au genre féminin, et le nombre singulier est signalé par l'absence de marque, c'est-à-dire par défaut (voir plus loin). Cela dit, on peut aussi considérer qu'au sein du genre I, /m-/ n'exprime que la classe nominale — et donc le singulier par défaut — tandis que /wa-/ exprime la cumulation [I-pluriel].

Plus remarquable est la flexion préfixale du swahili et des autres langues bantoues, dans la mesure où les langues connues montrent une nette préférence pour la suffixation. Selon Dryer (2013), dans un échantillon de 828 langues où l'affixation est assez développée pour que la statistique ait un sens, 529, soit environ 64%, sont suffixales, de façon

<sup>46</sup> Un auteur américain (Hofstadter 1985 : 159 et suiv.) a imaginé un anglais parallèle (et dystopique) qui, au lieu des pronoms masculin *he* et féminin *she*, aurait *whe* pour les humains blancs (*white*) et *ble* pour les humains noirs (*black*). Cette langue serait inhumaine dans tous les sens du mot.

<sup>47</sup> Les étymons latins *suffixus*, *praefixus* et *affixus* veulent dire littéralement « attaché dessous » (comprenez « après »), « attaché devant » et « attaché contre ».

<sup>48</sup> Ou de l'espèce entière (cf. chapitre 6).

<sup>49</sup> A la différence du genre grammatical des langues indo-européennes et sémitiques, le genre grammatical des langues bantoues se fonde non sur la division sexuelle, mais sur une classification des entités naturelles (humains, animaux, végétaux, artéfacts, etc.). Classification qui, de même qu'en français ou en arabe, n'a rien de systématique, si bien que le genre de bien des lexèmes est de fait arbitraire.

prédominante ou majoritaire. Les raisons de ce déséquilibre sont mal comprises. On a avancé l'hypothèse que les préfixes, parce qu'on les perçoit avant la base qui communique le sens du lexème, retardent la reconnaissance de celui-ci. L'étude de Gwilliams et al. (2022) décrite plus haut semble aller dans ce sens. Reste que 152 langues, soit environ 18,50% de l'échantillon, sont principalement ou majoritairement préfixales, et 147, soit 17,75%, font un usage à peu près égal des deux procédés. On relativisera ces résultats en tenant compte de ce que 828 langues sur un total de 6000, dont la plupart peu ou pas décrites, constituent un échantillon bien petit — sans oublier les milliers et milliers de langues qui se sont parlées depuis que les humains bavardent, voilà peut-être 300 000 ans, voire davantage, et qui ont disparu sans laisser de traces.

Beaucoup plus rare que la suffixation et la préfixation, l'infixation se place à l'intérieur de la base. Le jeu de langage dit « javanais » en offre un exemple en français, qui consiste à insérer /av/ entre l'attaque consonantique et le noyau vocalique de chaque syllabe d'un mot : *gros* 'grosse' > *gravos* 'gravosse' *bōzur* 'bonjour' > *bavōzavur* 'bavonjavour'. La séquence /-av-/ n'est qu'un brouillage, elle n'a aucun sens. Par contre, en tagalog, langue austronésienne parlée aux Philippines, /-um-/, infixé à une base verbale dans les mêmes conditions que /-av-/, exprime, entre autres choses, une action accomplie : p.ex. *sulat* 'écrire' vs. *s-um-ulat* 'avoir écrit'.<sup>50</sup>

Les affixes externes sont susceptibles de s'attacher les uns aux autres, presque toujours dans un ordre déterminé. Ainsi, le futur ukrainien *čytatymemo* /ʃi'ta-ti-m-emo/ 'nous lirons' comporte trois suffixes après la base /ʃi'ta/ : /ti/ pour l'infinitif, /m/ pour le futur, /emo/ pour la 1<sup>ère</sup> personne du pluriel. Le swahili *nilisoma* /ni-li-soma/ 'j'ai lu' présente de même deux préfixes avant la base /soma/ : /ni/ pour la 1<sup>ère</sup> personne du singulier, /li/ pour le passé. Je ne connais pas d'exemple d'infices en chaîne.

Liée à l'affixation (suffixation, préfixation, etc.) est la distinction traditionnelle entre fusion et agglutination fondée sur la plus ou moins grande séparation des valeurs ajoutées à la base. On qualifie ainsi les formes hongroises *házat* /'ha:z-ɔt/ et *házakat* /'ha:z-ɔk-ɔt/ d'agglutinantes parce que les valeurs [pluriel] exprimé par /ɔk/ et [accusatif] (objet direct) exprimé par /ɔt/ ajoutées à la base /ha:z/ 'maison' sont bien distinctes l'une de l'autre.<sup>51</sup> La forme latine *domūs* ou *domos* est dite fusionnelle, en revanche, parce que le suffixe /ūs/ ou /os/ ajouté à la base /dom/ 'maison' cumule ces mêmes valeurs. La phonologie intervient également : *čytatymemo* est sans doute (peut-être) moins immédiatement analysable que le *nilisoma* du swahili, autre langue réputée agglutinante.

En réalité, hongrois, latin, swahili et ukrainien partagent le même type flexionnel, appelé « concaténatif » puisqu'il consiste à enchaîner des éléments (cf. Bickels & Nichols 2013a). La qualité qui les différencie est ce que les linguistes anglophones nomment *exponence*, dérivé de *exponent* 'exposant', terme adopté par les morphologues au sens particulier de manifestation sensible (« exposition ») d'une signification (cf. Bickels & Nichols 2013b). Je traduis donc par « exposance ».<sup>52</sup> On distingue ainsi les affixes monoexposants (*monoexponential*) ou, plus simplement, univoques, des affixes polyexposants (*polyexponential*) ou, plus simplement, cumulatifs. Les langues traditionnellement dites fusionnelles comme le latin sont celles dont les affixes tendent à la cumulativité ; celles dites agglutinantes comme le hongrois et le swahili tendent à l'univocité.

<sup>50</sup> Autres types de « fixes » : le circonfixe qui encadre la base comme dans les participes passés allemands comme *gesagt* /gə-zag-t/ 'dit', en sorte qu'aussi bien \*/gə-zag/ que \*/zag-t/ sont mal formés ; l'interfixe, sans valeur sémantique, qui soude les deux éléments d'un composé à la façon du /o/ de *frank-o#alāmā* 'franco-allemand'. Nous en rencontrons d'autres.

<sup>51</sup> Plus exactement, /k/ pour le pluriel, /t/ pour l'accusatif. La voyelle est une voyelle de liaison qui varie selon le vocalisme de la base : cf. *embereket* /'ember-ɛk-ɛt/ 'hommes (accusatif)' (cf. Sauvageot 1982).

<sup>52</sup> Pardon pour le néologisme, mais il existe déjà 312 mots formés sur ce modèle, toujours productif : cf. *abondance*, *bienpensance*, *résistance*, etc.

Je dis bien « tendent », car les langues sont rarement cohérentes en matière d'exposance. Comme celui de l'occitan *lausetas* ou le <s> orthographique du FO <alouettes>, le /s/ du portugais *cotovias* 'alouettes' est univoque, il n'exprime que la pluralité.<sup>53</sup> Par contre, la finale /ɛj/ de *cantei* '(je) chantai' dans la même langue est cumulative, exprimant à la fois le temps (prétérit alias passé simple) — comparez *canto* '(je) chante' — et la 1<sup>ère</sup> personne du singulier — comparez *cantaste* '(tu) chantas'. (Pour *canto*, la notion de défaut permet d'arguer — mais je n'affirme rien — que /u/ <o> n'exprime que la 1<sup>ère</sup> personne du singulier, tandis qu'on peut supposer l'existence de raisons sémantiques pour considérer le présent comme le temps qui s'impose si aucun autre n'est explicitement signalé, en sorte que /u/ <o> serait univoque. En tout état de cause, ce recours à la sémantique paraît assez improbable s'agissant du prétérit.) Même les langues qui, tel le swahili, s'approchent de la complète univocité ne sont pas exemptes de cumulations : le correspondant négatif de *nilisoma* est *sikusoma* 'je n'ai pas lu', où /si/ cumule 1<sup>ère</sup> personne du singulier et négation, /ku/ passé et négation — si bien que la négation est exprimée deux fois, sans s'annuler comme en algèbre, autre écart de l'idéal univoque.

Le second procédé flexionnel synthétique, on l'appellera flexion non-linéaire (ou non-concaténative) parce qu'il n'opère pas sur la ligne des phonèmes segmentaux — ainsi nommés parce qu'on peut les séparer les uns des autres. L'exemple le plus clair en est donné par les verbes dits « forts » des langues germaniques, ainsi en anglais *I sing* 'je chante' vs. *I sang* 'j'ai chanté'. Rien n'est ajouté à la base, c'est elle qui est modifiée, ici quant au timbre de sa voyelle, /æ/ au lieu de /i/. Dans *I bring* 'j'apporte' vs. *I brought* 'j'ai apporté', la modification est plus importante, la base /brɪŋ/ changée en /brɔː/ (/brɔːx/ en vieil anglais), le /t/ final un suffixe apparenté au /(ɪ)d/ des verbes « faibles », qui montre au passage que flexion concaténative et flexion non-linéaire sont compatibles au sein d'une même forme. Le terme technique pour cette sorte de modification est « apophonie » ou bien l'emprunt allemand *ablaut*. Le FP n'en montre pas de cas bien nets. L'alternance vocalique dans les verbes tels que *ʒət-e* 'jeter', *ʒə=ʒet* 'je jette', *nu=ʒət-ɔ̃* 'nous jetons' provient historiquement d'un phénomène phonologique, selon que la base portait ou non l'accent tonique (cf. Moignet 1988 : 51 et suiv.). Toutefois, dans la mesure où ce conditionnement s'est opacifié avec le temps, il est permis de considérer que les verbes de ce type se conjuguent sur deux bases d'emploi désormais imprédictible.

Les langues sémitiques (amharique, arabe, hébreu, etc.) offrent des exemples encore plus évidents de cohabitation des deux procédés flexionnels synthétiques. Soit les deux formes de l'arabe classique *ʔaktub* 'j'écris' et *katabtu* 'j'ai écrit'. Il est généralement admis que la base — la racine, pour employer le terme traditionnel sur lequel je reviendrai — y consiste en la suite ordonnée des trois consonnes /ktb/, tandis que la ou les voyelles insérées entre elles expriment des significations flexionnelles, ici l'aspect (inaccompli vs. accompli). On peut dès lors considérer les deux /a/ de *katabtu* comme un affixe discontinu baptisé « transfixe ».<sup>54</sup> C'est pourquoi les spécialistes n'utilisent pas d'ordinaire « apophonie » pour décrire le phénomène. Une autre analyse est toutefois praticable, à savoir qu'on a affaire à deux bases /ktub/ et /katab/ en relation apophonique à l'instar de *sing* et *sang*.<sup>55</sup> A l'inaccompli, la valeur 1<sup>ère</sup> personne du singulier est exprimée par le préfixe /ʔa-/ ; à l'accompli, par le suffixe /-tu/. La réalité du transfixe ne s'en fonde pas moins sur des faits empiriques, ainsi la formation de certains pluriels en hausa, langue tchadique (afro-asiatique,

<sup>53</sup> Qu'on n'objecte pas que /s/ exprime la 2<sup>e</sup> personne du singulier dans *cantas* '(tu) chantes'. Il ne s'agit pas du même suffixe, mais bien d'un suffixe différent, quoique homophone.

<sup>54</sup> Circonfixe et transfixe sont tous deux discontinus. Mais le premier encadre la base, tandis que le second s'y intercale au moins en partie.

<sup>55</sup> A un niveau un peu plus abstrait, on peut considérer les deux bases comme étant /ktub/ et /ktab/. L'exclusion des groupes consonantiques initiaux en arabe classique explique l'insertion de /a/ entre /k/ et /t/ dans *katabtu*. Dans *ʔaktub*, le groupe n'est pas initial et peut donc se maintenir.

donc lointainement apparentée à l'arabe) : cf. *gulbi* / *gulàbe* 'ruisseau(x)', le singulier apparemment inanalysable, le pluriel comportant un transfixe /... à... e/, d'où /gul-à-b-e/ (cf. Caron 2015 ; Guzmán & Becker 2017).<sup>56</sup>

L'apophonie est à distinguer de la métaphonie (alias inflexion vocalique, alias umlaut) telle qu'elle se manifeste, par exemple, dans le yiddish *fus* 'pied' vs. *fis* 'pieds'. La distinction est avant tout historique. Dans les langues indo-européennes, les variations apophoniques semblent présentes dès le stade reconstitué dit « proto-indo-européen », correspondant à un ensemble de langues et/ou de dialectes parlés il y a quelques 5000 ans sur le territoire de l'actuelle Ukraine. On les explique, de façon hypothétique, par l'influence de consonnes « laryngales » disparues de toutes les langues subséquentes à l'exception du hittite (cf. Martinet 1986). La métaphonie est plus récente et observable surtout dans les langues germaniques où la modification de la voyelle de la base — phonologiquement un relèvement : /a/ → /æ/, /o/ → /ø/, /u/ → /y/ — est due à l'action à distance d'une voyelle haute suffixale (assimilation régressive). Celle-ci a disparu de l'exemple yiddish ci-dessus de même que du couple anglais *foot* / *feet* 'pied(s)' ; elle s'est maintenue dans l'équivalent allemand *Fuss* (/fus/) / *Füsse* (/fys-ə/) qui, compte tenu de ce que le /ə/ de l'allemand moderne remonte à un /i/, fait voir le mécanisme. Lequel est toutefois invisible dans *Vater* (/fa:tər/) / *Väter* (/fæ:tər/) 'père(s)' et maints autres exemples, si bien que la métaphonie est devenue un procédé en soi, synchroniquement indistinct de l'apophonie, et *Füsse* représente un autre cas de cohabitation des flexions concaténative et non-linéaire.

L'apophonie ne touche-t-elle que les voyelles ? La flexion non-linéaire ne peut-elle procéder par modification des consonnes ? Si, bien sûr. Les mutations consonantiques des langues celtes l'illustrent à merveille, par exemple la déclinaison irlandaise : cf. nominatif-accusatif singulier *an cat* /ə=kət/ 'le chat', nominatif-accusatif pluriel *na cait* /nə=kət/ 'les chats', génitif singulier *bia an chait* /b'ja ə=xat/ 'la nourriture du chat', génitif pluriel *bia na gcat* /b'ja nə=gat/ 'la nourriture des chats' (The Christian Brothers 1980 : 38). Le nominatif-accusatif singulier représente la forme de base. Au nominatif-accusatif pluriel et au génitif singulier, la consonne finale est palatalisée, elle devient « mince » (*slender*) selon la terminologie traditionnelle. Au génitif singulier, la consonne initiale est en outre mutée d'occlusive à spirante, c'est la mutation dite « lénition » (*lenition*) ou « adoucissement ». Au génitif pluriel, la consonne initiale est voisée (*eclipsis*).

En peul, langue Niger-Congo (groupe atlantique) parlée en Afrique de l'Ouest, l'opposition de nombre se marque également par un système de mutation des consonnes initiales : cf. *gorko* / *worbe* 'homme(s), mari(s)' où /ko/ et /be/ sont des marques de classe nominale, respectivement singulière et plurielle (cf. Noyé 1974 : 103-111). En albanais, pour une certaine classe de noms, la consonne finale se palatalise au pluriel : cf. *zog* /zɔg/ vs. *zobj* /zɔj/ 'oiseau(x)', *bujk* /bujk/ vs. *bujq* /bujc/ 'cultivateur(s)' (Boissin 1975 : 65).

Affectant les segments (voyelles et consonnes) la flexion non-linéaire doit a fortiori concerner les éléments dits « suprasegmentaux » parce qu'ils se superposent aux segments en se liant aux sommets de syllabe, soit le plus souvent aux voyelles. Ce sont en particulier l'accent tonique (*stress*), dont le principal corrélat acoustique est l'intensité, et les tons, c'est-à-dire la hauteur de la voix. Quant au premier, il y a peu d'exemples que sa place dans le mot joue un rôle flexionnel, alors que son rôle lexical — distinguer des lexèmes par ailleurs formellement identiques — est bien connu et documenté : cf. italien *ancora* /an'kora/ 'encore' vs. *ancora* /'ankora/ 'ancre', anglais *record* /'rɛkɔ:d/ 'compte-rendu, registre, disque' vs. *record* /rɪ'kɔ:d/ 'enregistrer'. En espagnol et en italien, avec les verbes réguliers, la prosodie accentuelle fait seule la différence entre la 1<sup>ère</sup> personne du singulier du présent de l'indicatif

<sup>56</sup> La formation du pluriel en hausa, comme généralement dans les langues afro-asiatiques, est notoirement complexe, quoique pas tout à fait imprédictible (voir la seconde étude citée). L'accent grave note le ton bas. Les voyelles non accentuées sont au ton haut.

et la 3<sup>e</sup> personne du singulier du prétérit : cf. *canto* /'kanto/ 'je chante' vs. *cantó* et *cantò* /kan'to/ 'elle/il chanta'.<sup>57</sup> En arabe de Juba et en kinubi, créoles arabes parlés l'un au Soudan du Sud, l'autre au Kenya et en Ouganda, un certain nombre de verbes forment le passif par un déplacement d'accent tonique de la première à la dernière syllabe : p.ex. *kátulu* 'tuer' vs. *katulú* 'être tué' ; et quelques noms se pluralisent par le même procédé : p.ex. *bágara* 'vache' vs. *bagará* 'vaches' (cf. Manfredi & Petrollino 2013 ; Lufin 2013 ; et cf. Kihm 2018 pour l'origine de ces formes). Une recherche plus poussée révélerait sans doute d'autres exemples.

Nombreuses, en revanche, sont les langues où des contrastes tonals remplissent une fonction flexionnelle.<sup>58</sup> Je n'en veux qu'un exemple, des plus clairs, le chinantec de Tlapeuzco, langue otomangue parlée au sud du Mexique (cf. Palancar & Léonard 2016). Les verbes s'y fléchissent à trois aspects, accompli, inaccompli et irréel, et quatre personnes : 1<sup>ère</sup> du singulier (1SG), 1<sup>ère</sup> du pluriel (1PL), 2<sup>e</sup> et 3<sup>e</sup> sans distinction de nombre (2, 3). Il y a trois tons fixes : haut (H), bas (B) et moyen (M) et des tons modulés, par exemple montant bas-haut (BH) ou montant bas-moyen (BM). Avec cela, la flexion du lexème HÚʔ 'courber', telle que donnée par les auteurs, est *húʔ<sup>B</sup>* à l'accompli 1SG ou 2 ; *húʔ<sup>BH</sup>* à l'accompli 1PL ou l'irréel 1SG, 1PL et 2 ; *húʔ<sup>M</sup>* aux trois aspects à la personne 3 ; *húʔ<sup>BM</sup>* à l'inaccompli 1 et 2. Le système apparaîtra plus clairement avec le tableau ci-dessous :

	1SG	1PL	2	3
accompli	B	BH	B	M
inaccompli	BM	BM	BM	M
irréel	BH	BH	BH	M

Certaines régularité en ressortent. La personne 3 est toujours au ton moyen ; à l'inaccompli et à l'irréel toutes les personnes sauf la troisième portent le même ton modulé, BM pour l'inaccompli, BH pour l'irréel ; les personnes 1SG et 2 distinguent tonalement les trois aspects. On relève des faits du même ordre dans d'autres langues d'Afrique, d'Amérique et d'Asie. En Europe, les « accents » du grec ancien — en fait des tons — remplissent des fonctions flexionnelles analogues. Ainsi, le nominatif et le datif singuliers du mot voulant dire « tête » ne se distinguent que par le ton, aigu (sans doute montant) pour le premier, circonflexe (sans doute modulé) pour le second : κεφαλή /kep<sup>h</sup>alé:/ vs. κεφαλῆ /kep<sup>h</sup>alê:/ (cf. Ragon 1962).<sup>59</sup>

Rien n'interdit de considérer alors le ton ou tonème — et l'accent tonique dans les exemples d'arabe de Juba et de kinubi — comme une marque flexionnelle analogue à un affixe, ce pour quoi on a forgé le terme de « suprafixe » et ajouté des « lettres tonales » à l'API, p.ex. ʔ pour le ton le plus haut, ʔ pour le ton moyen, etc, si bien que *húʔ<sup>M</sup>* s'écrit aussi *húʔʔ*.<sup>60</sup> Un affixe particulier, néanmoins, puisque sa matérialité phonique fait qu'il ne peut se réaliser que simultanément à l'articulation d'une voyelle — d'où le terme plus rarement employé de « simulfixe ». Cette propriété le rapproche de l'apophonie. Car, de même que pour celle-ci, l'émergence de tons dans une langue (la « tonogenèse »), se révèle souvent, pourvu qu'on puisse remonter assez loin dans le passé, être due à un contact au niveau segmental (cf. Michaud & Sands 2020). On observe ainsi qu'une consonne voisée tend à abaisser la hauteur de la voyelle qui la suit. Dès lors, si cette consonne disparaît, l'éventuelle

<sup>57</sup> Le galicien *cantou* /kan'tov/ a conservé la diphtongue responsable de l'accentuation finale en espagnol.

<sup>58</sup> Bien entendu, les tons, souvent appelés tonèmes par analogie avec les phonèmes, ont aussi, plus souvent en fait, une fonction lexicale.

<sup>59</sup> Le « iota souscrit » de κεφαλῆ est purement orthographique. Il y a aussi le lituanien, mais je recule devant la complexité des données.

<sup>60</sup> En arabe de Juba et kinubi, l'accent tonique initial représente le cas par défaut, commandé par la structure syllabique. C'est l'accent tonique final qui a valeur flexionnelle. Un mot ne pouvant comporter qu'un seul accent tonique, son attache entraîne automatiquement la disparition du concurrent.

différence de hauteur par rapport aux voyelles environnantes deviendra une marque en soi, de la même façon que l'alternance /u/ vs. /i/ dans le yiddish *fus / fis* 'pied(s)'.

Puisque la flexion non-linéaire repose sur la concurrence de bases phonologiquement différentes, le supplétisme, c'est-à-dire la mise en œuvre de bases *entièrement* différentes pour la flexion d'un même lexème, en représente le cas limite : cf. *ale* 'aller',  $\mathfrak{z}\partial=ve$  'je vais',  $\mathfrak{z}=i-r\epsilon$  'j'irai' ; ou encore  $\alpha j$  'œil /  $j\theta$  'yeux'. L'origine la plus commune du supplétisme est que des formes prises à plusieurs verbes se voient réunies au sein du même paradigme, en sorte que leurs différences sémantiques initiales s'effacent. Ainsi, deux verbes latins, *eo* et *uado* ont concouru de façon assurée à la conjugaison d'ALLER ; l'étymologie des formes en /al-/ est par contre incertaine. La paire  $\alpha j$  'œil /  $j\theta$  'yeux' représente un cas plus rare, où les deux formes ont même origine, latin tardif ou proto-roman \**oclu* et \**oclos*, respectivement nominatif-accusatif singulier et pluriel d'OCLU (latin classique *oculus* 'œil'), mais ont phonétiquement divergé au point de n'être plus mutuellement identifiables. De même  $b\alpha f$  'bœuf' /  $b\theta$  'bœufs',  $\alpha f$  'œuf' /  $\theta$  'œufs', etc., quoique la différence soit alors moins extrême. Le supplétisme ne se rencontre pas dans toutes les langues, tant s'en faut, et, là où il existe, il ne touche jamais qu'un petit contingent de mots. Une langue dont toutes les formes seraient supplétives est imaginable, mais elle chargerait la mémoire à un point qui la rendrait sans doute difficile à acquérir. Qui plus est, chaque forme, ne ressemblant à aucune autre, tendrait à se muer en signal holistique, ce qui serait la négation de la double articulation, propriété fondamentale de la faculté de langage humaine.

Les deux procédés synthétiques, concaténation et non-linéarité, partagent une caractéristique : leur produit est un mot unique, indécomposable à la façon de l'anglais *sang* / $s\alpha\eta$ /, ou bien morphologiquement décomposable comme  $\mathfrak{f}\tilde{a}t-\epsilon$ , mais de telle sorte que les parties ainsi dégagées, base et affixe, n'ont aucune existence autonome, sauf s'il se trouve que la base coïncide avec un mot-forme — auquel cas il n'en reste pas moins nécessaire de séparer les deux niveaux. En d'autres termes, mot-forme est un concept de probable réalité cognitive ; les concepts de base et d'affixe, par contre, n'ont sans doute de réalité que théorique. Le premier appartient au langage et aux langues, les seconds aux théories qui tâchent d'en expliquer le fonctionnement. Il est permis d'en conclure que concaténation et non-linéarité ne représentent pas deux types flexionnels distincts, mais deux modes de réalisation de la flexion synthétique. On a vu avec les pluriels métaphoniques de l'anglais et du yiddish qu'il n'y a rien d'exceptionnel à ce qu'un mode se transforme en l'autre au cours de l'histoire (cf. Bybee & Newman 1995).

L'autre type de flexion, l'analytique, ne partage pas cette caractéristique d'unicité. Tout d'abord, entre l'autonomie et son absence il existe un tiers, dont l'article défini français est un bon exemple. Soit l'expression définie  $l\partial=gro\_fa$  'le gros chat'. La propriété de définitude que l'article exprime, il l'attribue certes au SN [gro fa], mais en dernière analyse au lexème CHAT et à son référent, dont la corpuence n'est qu'une qualification incidente. Par là, l'article défini ne fonctionne pas autrement qu'un affixe de nombre, ce que confirment les langues où il est en effet affixal, par exemple les langues scandinaves et sémitiques. Pourtant, en français, en roumain et ailleurs, il se distingue des affixes en cela qu'il ne s'attache pas nécessairement à ce qu'il module. L'exemple le montre :  $l\partial$  est lié à *gro*, soit plus généralement au premier constituant du SN dont rien d'extérieur au SN ne peut en principe le séparer, quoiqu'il faille ne jurer de rien en ce domaine — cf. ?  $l\partial$ ,  $\mathfrak{z}\partial$  *krwa*,  $gro\_fa$  ? 'le, je crois, gros chat', que je ne me risquerai pas à déclarer impossible. Mais il se distingue des mots autonomes en cela qu'il *doit* (sauf exception) s'attacher à quelque chose. Les éléments qui présentent la conjonction de ces deux propriétés sont, on les aura reconnus, les clitiques décrits à la section précédente. La flexion procède donc aussi par le biais d'attachement de clitiques. De ce fait, elle cesse d'être strictement synthétique, pour autant que  $l\partial=fa$  'le chat' n'est pas un mot unique ou, mieux dit, compact, au sens qu'à l'adjectif en parler ordinaire —

« dont les parties sont étroitement serrées et ne se séparent que difficilement », dit Larousse — voire (mais je ne m'avance pas trop) en topologie.<sup>61</sup> J'adopterai pour la définir le terme de « semi-analytique » proposé par Creissels (2006a : 165-166).

Les exemples les plus clairs de flexion semi-analytique nous viennent des langues créoles. On notera d'abord que, s'agissant du marquage de la définitude, l'haïtien est comme l'image en miroir du français, le déterminant défini se lie au dernier élément du SN et il est enclitique : cf. *chat ki pa renmen dlo a* {chat qui NEG aimer eau DEF} 'le chat qui n'aime pas l'eau'. *Dlo* 'eau' étant générique et la généricité toujours signifiée en haïtien par un nom sans déterminant, on est assuré que la définitude exprimée par *a* porte sur la tête du SN, *chat* 'chat'. Mais que le déterminant soit morphologiquement attaché à ce qui le précède immédiatement est d'autre part garanti par la forme /a/, forme alternative (allomorphe) de l'article suscitée par tout mot que termine une voyelle orale : comparez *pen an* /pẽ=ã/ 'le pain', *chat la* /ʃat=la/ 'le chat', *fanm nan* /fã=nã/ 'la femme'.

Encore plus exemplaires sont les systèmes verbaux. Les créoles n'ont pratiquement rien conservé des flexions verbales synthétiques de leurs langues sources (anglais, arabe, espagnol, français, néerlandais, portugais), et les ont remplacées par des systèmes de « particules » (aussi appelées « marqueurs ») préposées. Le martiniquais est typique de cette évolution. Moyennant une bonne dose de simplification, on peut dire qu'il distingue quatre temps, présent, prétérit, antérieur et futur, soit *Man ka vini* {je KA venir} 'Je viens', *Man vini* {je venir} 'Je vins / suis venu', *Man té vini* {je TE venir} 'J'étais venu', *Man ké vini* {je KE venir} 'Je viendrai', *Man té ka vini* {je TE KA venir} 'J'étais en train de venir'.<sup>62</sup> A l'évidence, ces cinq formes (et d'autres encore) constituent le paradigme flexionnel du lexème verbal VINI tout autant que *vjẽ* 'viens | vient', *vãne* 'venais | venait', etc. celui de VENIR pour les propriétés dites de temps-aspect-mode (TAM). La différence est typologique : flexion entièrement semi-analytique plutôt que majoritairement synthétique, avec pour corrélat une complète transparence du système. Connaître le verbe, forme et sens, et les trois proclitiques flexionnels suffit pour connaître tout le paradigme.<sup>63</sup>

Les particules *ka*, *ké* et *té* sont toujours rigoureusement contiguës au verbe qu'elles modifient.<sup>64</sup> On pourrait donc aussi bien y voir des préfixes, d'autant qu'aucune des quatre conditions discriminantes de la section précédente ne trouve à s'appliquer. L'usage, sanctionné par la graphie, est plutôt d'y voir des proclitiques, dont l'adjacence à l'hôte justifie la distinction entre clitiques de mots et clitiques syntagmatiques. Y invite aussi le fait que d'autres créoles présentant des systèmes analogues, par exemple l'haïtien, tolèrent l'insertion d'adverbes modificateurs du prédicat tels que « souvent » entre le clitique et le verbe.

Enfin, faisant un pas de plus vers l'analyticité, on peut fléchir par la réunion de plusieurs mots autonomes, c'est ce qu'on appelle une *périphrase*, par exemple *ʒ=ɛ bjẽ dɔrmi* 'j'ai bien dormi'. Le passé dit composé appartient sans conteste au paradigme verbal, *ʒ=ɛ dɔrmi* 'j'ai dormi' l'une des formes qui réalisent le lexème DORMIR, si ce n'est que ce n'est pas un mot, mais bien une alliance, une collocation de deux mots autonomes, *avwar* 'avoir' fléchi au présent de l'indicatif, suivi du participe passé — mais pas forcément contigu à lui

<sup>61</sup> Un espace topologique est compact s'il est séparé, c'est-à-dire si deux points quelconques y peuvent toujours appartenir à des voisinages disjoints. Quoiqu'il faille se méfier des analogies faciles, surtout avec des domaines dont on n'est pas spécialiste, il semble bien que les lexèmes, dont les mots-formes constituent le voisinage, soient dans ce cas si on les conçoit comme des points dans l'espace lexical. C'est donc celui-ci qui serait compact au sens mathématique du terme, plutôt que les lexèmes et les mots-formes.

<sup>62</sup> *Vini* 'venir' est un verbe dynamique dénotant une activité. Avec les verbes dénotant un état (statifs), par exemple *enmen* 'aimer', l'absence de particule produit un présent (*I enmen mwen* 'Elle/Il m'aime'), et *té* un passé (*I té enmen mwen* 'Elle/Il m'a aimé / m'aimait'). Sans doute pour des raisons sémantiques, la combinaison *té ka* est rare, voire inexistante avec les verbes statifs.

<sup>63</sup> On nomme « entropie conditionnelle » la mesure de cette prédictibilité des formes les unes à partir des autres.

<sup>64</sup> *Té* /te/ provient à l'évidence du FP (dix-septième siècle) *ete* 'étais | était | étaient'. L'étymologie de *ka* et *ké* est obscure.

comme le montre l'exemple. La définition suivante est une version adaptée de celle proposée par Ackerman et al. (2011 : 331 ; voir aussi Bonami 2015). Soit un lexème appartenant à une certaine partie du discours (p.ex. DORMIR, *verbe*) et un ensemble d'une ou plusieurs propriétés appropriées (p.ex. {PERSONNE-NOMBRE 3SG, TEMPS *passé*, MODE *indicatif*}). On qualifie de périphrastique l'association du lexème et des propriétés si celle-ci s'effectue par la cooccurrence (collocation) de l'un des mots-formes manifestant le lexème (p.ex. *dormi*) et d'un ou plusieurs mot-formes (p.ex. *ε*) manifestant un autre lexème qu'on qualifiera d'auxiliaire. Chacun de ces mots-formes constitue alors la tête d'un syntagme, p.ex. *ε* du syntagme auxiliaire (SAux) et *dormi* du syntagme verbal (SV). La réunion de ces syntagmes diffère selon les langues. Certaines requièrent que les têtes en soient adjacentes, ainsi le parfait persan : cf. *xaride ast* {acheter.PP être.3SG.PRS} 'elle/il a acheté' (Lazard 2006 : 138) ; pour d'autres comme le FP, elles peuvent être séparées par des composantes du syntagme modifié, par exemple un adverbe, ou bien par des expressions extérieures à la périphrase telles des incises : cf. l'exemple donné plus haut de  $\mathcal{Z}=\varepsilon$  *bjê dormi* 'J'ai bien dormi', ou encore  $\mathcal{Z}=\varepsilon$ , *i(l)=mə=sābl, dormi lītā* 'J'ai, il me semble, dormi longtemps'. On le verra au chapitre 9, le mode périphrastique joue un rôle important dans la flexion verbale du FP.

A ces procédés, enfin, il convient d'ajouter le *défaut*, à savoir la signification par absence de signifiant. On vient d'en voir un exemple avec le martiniquais *Man vini* 'Je suis venu'. En l'absence d'une marque clitique de TAM (voir ci-dessus), l'interprétation du prédicat dépend de la nature sémantique du verbe (voir note 42). Cette vision des choses ne fait toutefois pas l'objet d'un consensus parmi les linguistes. Beaucoup préfèrent considérer que, dans ce cas comme dans les autres, le verbe est bien associé à une marque, sauf que celle-ci, à la différence des autres, est dépourvue de substance phonétique, c'est, pour employer le terme consacré, une marque « zéro ». Les marques zéro ne se prononcent ni ne s'entendent, mais elles n'en ont pas moins un contenu sémantique, en l'occurrence et pour rendre compte de l'interprétation variable, une valeur aspectuelle perfective ou accomplie, partant du principe qu'une activité accomplie est terminée au moment où on la dit telle, tandis qu'un état accompli est en vigueur à ce même moment.

Le raison principale de cette hypothèse me paraît être l'attrait, par ailleurs bien légitime, pour la symétrie. J'en donne encore deux exemples. Soit l'anglais *the cat* 'le chat' vs. *the cats* 'les chats' : l'article anglais étant invariable, rien n'indique que le premier syntagme est singulier, alors que la pluralité du second ressort clairement du /s/ final de *cats*. Dès lors que ce dernier ne peut s'analyser que comme /kæt-s/, /-s/ voulant dire « pluriel », la tentation est forte d'analyser le premier comme /kæt-Ø/, en lui adjoignant un suffixe signifiant « singulier », qui ne diffère de /-s/ que du fait d'être phonologiquement vide.<sup>65</sup> Autre exemple : en finnois, un verbe à la 3<sup>e</sup> personne du singulier du présent ne montre aucune désinence, à la différence des autres personnes, si la base se termine par une voyelle longue ou une diphtongue : cf. *syö* 'elle/il mange' vs. *syön* 'je mange', *syöt* 'tu manges', *syömmе* 'nous mangeons', *syötte* 'vous mangez', *syövät* 'elles/ils mangent' (Sauvageot 1949 : 64). Là encore, analyser *syö* comme /syø-Ø/ paraît raisonnable (cf. Haspelmath & Sims 2010 : 64).<sup>66</sup>

Raisonné et séduisant, mais qui néglige une réalité essentielle : l'absence est par elle-même signifiante. Les systèmes sémiotiques autres que le langage en témoignent. En signalisation routière, il existe des panneaux de sens interdit et des panneaux de sens obligatoire, il n'en existe pas de sens autorisé. Cela signifie qu'en l'absence d'indication du

<sup>65</sup> Quoique moins répandus que les pluralisants, les affixes singulatifs se rencontrent, p. ex. en breton : *melc'hwed* 'des limaces' vs. *eur velc'hwedenn* 'une limace', *plouz* 'de la paille' vs. *eur blouzenn* 'une paille', *stered* 'des étoiles' vs. *eur steredenn* 'une étoile' (Trépos 1956). Dans ce cas, c'est le pluriel, ou plutôt le collectif qui est censé porter une marque zéro.

<sup>66</sup> Mais si la base se termine par une voyelle brève, celle-ci s'allonge : p.ex. *sanoo* /'sano:/ 'elle/il dit'. L'allongement peut alors être considéré comme une marque explicite.



contraire, *par défaut*, il est permis d'emprunter une voie ou de ne pas l'emprunter.<sup>67</sup> De même, en l'absence de /-s/ final, par défaut, *cat* est singulier ; en l'absence de désinence, un verbe finnois à radical long ou diphtongué renvoie à la 3<sup>e</sup> personne du singulier ; en l'absence de particule, par défaut, un verbe martiniquais se comprend comme aspectuellement accompli. On peut donc s'en tenir à la tautologie : quand il n'y a rien, il n'y a rien, et c'est ce rien qui signifie par son contraste avec quelque chose. Que la maîtrise de l'opposition privative — puisque c'est le terme technique pour cette relation entre une marque et son absence — soit un apanage de la raison humaine, on n'en peut guère douter.<sup>68</sup> Et puisqu'elle suffit à expliquer les phénomènes illustrés à l'instant, pourquoi aller supposer des entités dont la nécessité est rien moins qu'assurée, en contravention avec le principe de parcimonie, jamais mieux formulé que par Guillaume d'Ockham (1285-1347) : « *Entia non sunt multiplicanda praeter necessitatem* » 'Il ne faut pas multiplier les entités plus qu'il n'est nécessaire' ?

Mais n'est-ce pas chicaner ? Quelle différence entre rien et zéro ? La différence est que zéro n'est pas rien. Le zéro mathématique a des propriétés, le zéro linguistique en a aussi. On s'en est servi pour rendre compte des pluriels non-linéaires de l'anglais tels que *foot / feet* 'pied(s)', *mouse / mice* 'souris', etc. Ces noms formeraient leur pluriel par suffixation comme le fait la majorité, non pas d'un /-s/, mais d'un zéro imprégné du pouvoir de déclencher la modification métaphonique de la voyelle du radical (cf. Halle & Marantz 1993). Dans l'optique parcimonieuse ici préconisée, c'est la métaphonie elle-même qui signifie la pluralité, procédé minoritaire en anglais, dont l'explication est à chercher dans l'histoire de la langue, comme expliqué plus haut. Il est vrai que de telles conclusions ont quelque chose de décevant. N'y a-t-il rien à trouver derrière les apparences en matière de langage ? A cette question, on peut opposer celles-ci : la réalité est-elle tenue de toujours nous surprendre ? N'est-ce pas un travail déjà assez ardu que de la décrire telle qu'elle se présente ?

En tout cas, l'importance du défaut est caractéristique du FP — ce en quoi, il n'est peut-être pas sans intérêt de le remarquer, il continue la tradition de l'ancien français où les flexions « amorphes » étaient monnaie courante en flexion nominale aussi bien que verbale : p.ex. *arc* 'arc(s) (cas régime singulier / cas sujet pluriel)' vs. *ars* /ar-s/ 'arc(s) (cas sujet singulier / cas régime pluriel)' ; *ber* 'baron (cas sujet singulier)' vs. *baron(s)* 'baron(s) (autres cas)' ; *aim* 'j'aime' vs. *amons* 'nous aimons' (cf. Zink 1989). Ces exemples — dus à des changements phonétiques entre le latin tardif et l'ancien français — ont été régularisés en moyen français, mais de nouveaux changements phonétiques en ont produit d'autres.

Les propriétés — encore appelées « traits » (*features*) — associées par ces divers procédés synthétiques et (semi-)analytiques sont dites « morphosyntaxiques », car elles peuvent se manifester dans les deux dimensions, comme on le voit par l'ukrainien *koty spljat* 'les chats dorment' où la pluralité est marquée morphologiquement sur le nom (*kit* 'chat' vs. *koty* 'chats') et syntaxiquement, par l'accord du verbe à la 3<sup>e</sup> personne du pluriel (cf. *kit spyt* 'le chat dort'). (On comprendra mieux au chapitre 6 pourquoi je ne prends pas mon exemple au français.) On est ainsi amené à distinguer entre flexion inhérente et flexion contextuelle (cf. Kuryłowicz 1964 ; Booij 1996). Pour un nom (en ukrainien), référer à une ou plusieurs entités est une propriété inhérente, sans laquelle la signification est incomplète et la forme impossible.<sup>69</sup> Le verbe au pluriel, en revanche, ne dénote pas plusieurs occurrences de la même action, mais une seule action attribuée à plusieurs sujets. En FP, aucun accord ne

<sup>67</sup> Un système totalitaire parfait se signale par l'absence de défaut : tout ce qui n'y est pas explicitement interdit y est explicitement obligatoire. Heureusement pour l'humanité, il y a toujours un défaut. « *There's a crack in everything / That's where the light comes in* » (Leonard Cohen).

<sup>68</sup> Puisqu'il s'agit au fond de ne prêter attention qu'aux signaux qui sortent de l'ordinaire, il est probable que tous les animaux autres que les humains partagent cette capacité, qui favorise leur survie en leur évitant d'être en permanence dans un état de surveillance maximale qui inhiberait toute autre activité (se reproduire, manger...).

<sup>69</sup> Il importe de préciser de quelle langue on parle, vu l'existence de langues où le nombre n'est exprimé d'aucune façon, comme en chinois mandarin.

s'exprime avec les verbes réguliers : cf. *ẽ=fa mjol* 'un chat miaule' vs. *de=fa mjol* 'des chats miaulent'.

En principe — il y a comme toujours des exceptions, mais peu nombreuses — les propriétés flexionnelles s'appliquent à tout ce pour quoi elles ont pertinence : tout procès ou état peut concerner le présent (*i(l)=plø* 'il pleut'), le passé (*i(l)=pløve* 'il pleuvait') ou le futur (*i(l)=pløvra* 'il pleuvra') ; par conséquent, tout verbe dénotant un procès ou un état est susceptible de se fléchir pour le temps. (Cela se vérifie dans l'ensemble en français, FO comme FP.<sup>70</sup> Pas en chinois mandarin. Mais c'est que les langues varient beaucoup quant à ce qui, des possiblement exprimés, l'est obligatoirement.) C'est pourquoi on caractérise la flexion comme entièrement productive. La morphologie dérivationnelle, pour sa part, ne l'est en général que partiellement : *ty#ær* 'tueur' se dérive de *ty-e* 'tuer', mais non, du moins « régulièrement », *?bles#ær* ?'blesseur' de *bles-e* 'blesser'. En comparaison des autres langues romanes, le français, en particulier son registre normatif, est notoirement « coincé » en matière de dérivation et de néologie.

Les questions soulevées dans cette section, en particulier la flexion non-linéaire, font qu'une brève digression s'impose avant de passer à d'autres points. Reprenons l'exemple albanais *zog /zog/* vs. *zogj /zoj/* 'oiseau(x)'. La tentation est grande de supposer que l'occlusive palatale finale résulte de l'action d'une voyelle haute, disons un /i/, à sa suite, qui s'est effacée de la prononciation, mais n'en constitue pas moins la « vraie » marque du pluriel. (On n'a pas oublié, j'espère, la discussion du chapitre 3 sur les consonnes « latentes » du français et la formation du féminin.) En tant qu'explication historique, cela se peut, même si l'albanais moderne ne montre pas (plus ?) de pluriels en /i/.<sup>71</sup> Mais que dire des albanophones du vingt-et-unième siècle ? Est-il bien raisonnable de supposer qu'elles et ils se livrent à cette sorte d'analyse ? N'est-il pas plus probable que, dans leur esprit, /zog/ est la seule forme possible s'il est question d'un seul oiseau, /zoj/ s'il est question de plusieurs ? Après tout, les francophones acceptent sans sourciller d'avoir deux *jø* quand un cyclope n'a qu'un *æj*, de manger un *æf*, mais d'acheter une boîte de six *ø*. Et les locutrices et locuteurs du chinantec n'ont évidemment aucune idée des possibles consonnes disparues — il y a combien de siècles, voire millénaires ? — responsables des marquages tonals des formes verbales qu'elles et ils ont acquises telles quelles. D'où deux impératifs. D'abord, ne pas mêler diachronie et synchronie hors des périodes de changement en cours, lorsque l'intersection des deux dimensions se manifeste par de la variation au sein de la communauté linguistique (Labov).<sup>72</sup> Deuxièmement, si l'on prend au sérieux l'idée selon laquelle les langues sont (aussi) des objets mentaux, ne pas imaginer des processus ou des règles dont la réalité cognitive est sujette à caution, voire improbable.

## 5.5. Le morphème

On s'étonne peut-être que ce terme, dont la finale *-ème* indique qu'il désigne l'unité morphologique de même que « phonème » désigne l'unité phonologique, « sème » l'unité

<sup>70</sup> Moyennant l'existence de verbes défectifs dont certaines formes a priori possibles ne sont pas en usage et souvent indécisés : cf. *ʒə=fri* 'je fris', *nu= ?* 'nous ?'.

<sup>71</sup> En roumain, dans une paire telle que *carte /'karte/* vs. *cărți /'kerts/* 'livre(s)', le /i/ final, indubitable marque de pluriel, quoique fort réduit, n'est pas tout à fait inaudible (cf. Pană Dindelegan 2010). Albanais et roumain appartiennent à la fameuse aire linguistique balkanique. Les deux langues présentent ainsi certains parallèles que leur lointaine parenté — ce sont deux langues indo-européennes, le roumain roman, l'albanais isolé — ne suffit pas à expliquer.

<sup>72</sup> Comme on l'a vu, les faits de flexion non-linéaire proviennent le plus souvent de processus phonologiques dus à des contacts de phonèmes dont l'élément déclencheur s'est amuï après avoir produit son effet. On sait qu'il en fut ainsi des mutations celtiques. Une fois la cause phonologique disparue, l'effet devient purement morphologique, c'est-à-dire lié aux propriétés portées par le mot-forme. Comme il est impossible qu'elle se produise en même temps chez tous les locuteurs, cette disparition et le changement de statut qui s'ensuit peuvent prendre plusieurs générations, d'où des périodes de variation.

sémantique, etc., ne soit pas encore apparu dans cet exposé. Peut-on parler de morphologie sans définir ce qui est censé en constituer l'élément ultime ?

Revenons au *doukipudonktan* quenellien. J'ai noté (note 5) que *py* 'pue' introduit une incertitude quant au nombre d'éléments, précisément parce qu'il en dit plus que la seule signification du lexème PUER, à laquelle il ajoute les propriétés [présent], [indicatif] et, dans cette phrase, [3SG], mais qu'aucun exposant ne semble correspondre à cet ajout, puisqu'il ne se décompose qu'en deux phonèmes /p/ et /y/ dénués de significations. En FO, le <e> de <pue> ne saurait davantage être considéré comme la marque du présent et/ou de la 3<sup>e</sup> personne du singulier, car il se lit aussi dans <je puerai>, <je pue> et <tu pues>. <sup>73</sup> Conformément à l'embryon de typologie développé à la section précédente, on peut considérer que *py* se fléchit par défaut. En l'absence de toute exposance de propriétés morphosyntaxiques, le mot-forme coextensif de la base se voit associé par défaut [présent] pour la temporalité, [indicatif] pour le mode et [1SG | 2SG | 3] pour la personne et le nombre, ces trois valeurs virtuelles discriminées par un pronom sujet clitique (*ty=py* 'tu pues' vs. *i(l)=py* 'il(s) pue(nt)') ou la présence d'un SN sujet (*set=afēr py* 'Cette affaire pue').

Si *py* 'pue(s) / puent' paraît morphologiquement inanalysable, il n'en va pas de même, on le sait, de *py-ḥ* 'puons', qui comporte bien deux parties signifiantes, deux exposants, /py/ et /ḥ/. Or, ce sont ces exposants qu'il est de tradition d'appeler des *morphèmes*, c'est-à-dire des associations minimales (élémentaires) d'une forme et d'un sens, des signes minimaux en d'autres termes, conformément à la définition saussurienne du signe. Il semble toutefois qu'on ait besoin d'une distinction plus fine, entre morphème et *morphe* (cf. Matthews 1972 : 41-42). Pour le voir, il y faudra un élément un peu plus complexe que cette terminaison -ḥ. Soit donc le génitif pluriel *canum* du latin *canis* 'chien' (*latratus canum* 'l'aboiement des chiens'). Deux parties décomposent exhaustivement cette forme : /kan/, la base, et /um/ la désinence. Aller plus loin ne livre que des phonèmes sans signification (/k/, /a/, /n/, /u/, /m/). En tant qu'objet phonémique distinguable dans l'ensemble des mots-formes réalisant le lexème CANIS — par exemple par comparaison-opposition avec *can-ibus* 'chiens (datif-ablatif)' — /um/ est un morphe, une forme (μορφή). Cette forme exprime une signification complexe : « génitif pluriel d'un nom de classe flexionnelle 3I (3<sup>e</sup> déclinaison imparisyllabique) », et c'est alors, prenant en compte le contenu de l'objet phonémique, qu'on parle de morphème. <sup>74</sup> Mais la complexité de la signification exprimée oblige à prendre une décision. Ou bien on admet que la signification « génitif, etc. » correspond à un ensemble de propriétés morphosyntaxiques empaquetées dans le morphe /um/ et collectivement exprimées par lui, ce qu'on peut noter comme ceci, où la double flèche symbolise une relation de correspondance réciproque. <sup>75</sup>

/-um/ ↔ {[CAS *génitif*, NOMBRE *pluriel*, CLASSE 3I]}

Figure x

Ou bien, pour des raisons théoriques abordées à la prochaine section, on insiste sur la nécessité d'une correspondance biunivoque (un à un) entre formes et significations, morphes et morphèmes. On se voit alors forcé, pour qu'il y ait autant de ceux-là que de ceux-ci dans le cas qui nous occupe, de poser de ces morphèmes zéro dont il a été question plus haut, auxquels correspondent des morphes dépourvus de substance phonétique, des formes immatérielles en quelque sorte. Si l'on suppose (arbitrairement) que /um/ signifie « génitif »

<sup>73</sup> Il s'agit en fait de la voyelle dite « thématique » caractéristique des verbes du 1<sup>er</sup> groupe, un « morphème » (cf. Aronoff x), à savoir un élément purement morphologique, dépourvu de justifications sémantiques ou syntaxiques. Rien ne lui correspond en FP.

<sup>74</sup> On appelle « imparisyllabiques » les noms qui présentent une syllabe de plus au génitif qu'au nominatif singuliers : cf. *legis* 'de la loi' vs. *lex* 'loi'. *Canis* est exceptionnel.

<sup>75</sup> Le signe égal (=) ferait également l'affaire.

et rien de plus, la notation devient ceci, où deux zéros se concatènent (+) à l'unique morphème apparent :

$/-um/ \leftrightarrow [\text{CAS } \textit{génitif}] + \emptyset \leftrightarrow [\text{NOMBRE } \textit{pluriel}] + \emptyset \leftrightarrow [\text{CLASSE } 3I]$

### Figure x

Toute association d'une forme avec une signification ou une propriété morphosyntaxique dont il paraît possible de se former un concept distinct, qu'elle soit exposée ou non, constitue dès lors un morphème, si bien que *can-um* se compose de quatre morphèmes, y compris la partie invariante, la base /kan-/ (cf. figure x). A noter toutefois qu'il règne là une certaine confusion terminologique (cf. la discussion chez Fradin 2003 : 30-41). C'est qu'il existe un usage de « morphème » qui n'applique le terme qu'aux éléments à signification fonctionnelle telle que pluriel, génitif, etc., ceux que Martinet (1965) nomme « monèmes grammaticaux » par opposition aux « monèmes lexicaux ». Afin d'éviter au possible cette confusion, et n'ayant pas l'usage de la distinction morphème-morphe comme on le verra, je continuerai à n'employer que le terme « exposant » réservé aux éléments à signification fonctionnelle, si bien que /kan-/ n'est pas un exposant, c'est la réalisation du lexème CANIS dans le mot-forme *can-um*, la base (ou radical ou thème) de la forme (*stem* en anglais).<sup>76</sup> Le principal avantage du terme est de ne suggérer aucune limitation quant à l'exposance : affixes, procédés non-linéaires, rien, autant de façons possibles d'exprimer des significations morphosyntaxiques.

En fait, la signification « génitif pluriel » est généralement considérée comme une propriété unique globalement exprimée par /-um/ pour une partie des noms de la 3<sup>e</sup> déclinaison. Cela se justifie par l'absence en latin (comme dans bien d'autres langues) d'un exposant du cas qui ne soit pas en même temps un exposant du nombre. Ce n'en est pas moins à l'évidence une infraction au principe de biunivocité qui exigerait un zéro supplémentaire pour l'une ou l'autre propriété — laquelle au fait ? Mais si l'on en autorise une, de quel droit ne pas les autoriser toutes ? Pourquoi ne pas adopter l'analyse de la figure x ?

Selon celle-ci, *kan-um* 'des chiens' est un mot-forme, partie d'un ensemble de paradigmes, c.-à-d. de listes de mots-formes qui sont autant de variations (flexions) du lexème CANIS. C'est cette appartenance et les oppositions qu'elle suscite qui font que la finale /-um/ s'individualise, devient perceptible et généralisable et mérite le tiret qui l'introduit. Ainsi, au sein du paradigme du pluriel, *can-um* s'oppose en cas au nominatif *can-es* et au datif-ablatif *can-ibus* ; il s'oppose en nombre au génitif singulier *can-is* 'du chien' ; il relève (exceptionnellement) de la classe 3I par opposition au parisyllabique (3P) *ciuis* 'citoyen' dont le génitif pluriel est *ciu-ium* 'des citoyens'. Nul besoin de morphèmes zéro ni, par conséquent, de la distinction entre morphème et morphe, puisque les propriétés morphosyntaxiques sont ou bien exposées, donc « morphiques », ou bien inférées par défaut de l'absence de forme. Dès lors, la notion d'exposant suffit et l'analysabilité des formes complexes persiste, c'est seulement que les exposants n'ont d'existence qu'au sein des paradigmes qui en permettent

<sup>76</sup> Un terme qui revient fréquemment dans les études de morphologie est « racine » (*root*). Je ne l'utilise pas, car la distinction entre racine et base n'a guère d'utilité heuristique en FP. On définit d'ordinaire la racine d'un lexème comme ce qui reste une fois ôtés les éléments flexionnels et les morphèmes (cf. note 53). Ainsi, /domin/ est la racine du latin *dominus* 'maître', dont /dominu/ est la base sur laquelle sont formés le nominatif /dominu-s/ et l'accusatif /dominu-m/, /u/ la voyelle thématique caractéristique de la 2<sup>e</sup> déclinaison. Là, racine et base sont distinctes et également abstraites, puisque inutilisables telles quelles. Avec les noms athématiques comme *canis*, *canem*, etc. 'chien', en revanche, l'abstraction persiste, mais racine et base ne se distinguent plus, toutes deux /kan/. En FP, comme le montre *py-ḫ*, l'ablation de la flexion /ḫ/ laisse quelque chose qui, non content d'être autant une racine qu'une base, est aussi un mot-forme : *(ty=)py* 'tu pues'. C'est là un phénomène très général, si bien qu'on peut en effet se passer de la distinction racine-base en FP et ne garder que le second terme, dont on appréciera l'indiscutable utilité descriptive. Comme le disait Saussure (1915/1982 : 257) : « Il ne faut donc pas s'étonner si le sentiment de la racine est fort peu développé en français. »

l'identification. En d'autres termes, les exposants sont *abstrait*s des mots qui les contiennent et se voient attribuer une signification par rapport à ces mots et à leurs paradigmes.

Dans cette hypothèse, on évitera donc de dire que *can-um* consiste en la composition de /kan-/ avec /-um/ (et plus) comme impliqué par le schéma de la figure x ; mais plutôt que, dans le mot-forme *canum*, la signification lexicale du lexème CANIS (« canis lupus familiaris ») est associée à la sous-séquence phonétique /kan/ (la base), et la signification complexe [CAS *génitif*, NOMBRE *pluriel*, CLASSE 3I] à la sous-séquence /um/, l'une et l'autre conjuguées au sein du mot-forme *canum* référant à l'état de fait que décrit l'énoncé « une pluralité d'entités animées de l'espèce *canis lupus familiaris* est décrite comme étant, entre autres possibilités, l'auteur d'une activité décrite par le terme nominal « aboiement » (*latratus*) ». D'où le schéma ci-dessous, où la relation pertinente n'est pas la concaténation (+) comme dans la figure x, mais la précedence linéaire (<):

{kanum ↔ {/kan/ ↔ CANIS < /um/ ↔ [CAS *génitif*, NOMBRE *pluriel*, CLASSE 3I]}}

Figure x

Les accolades extérieures indiquent que la signification complexe enclose dans les accolades intérieures est une propriété du mot-forme, au seul niveau duquel elle se voit décodée.

Que les sous-séquences dépendent pour leur signification du mot-forme qui les contient a pour conséquence logique qu'un même objet phonétique peut exprimer des traits différents, en tout ou partie, dans des mots-formes différents. Ainsi, en FP, au futur, /ð/ exprime la 1<sup>ère</sup> personne du pluriel dans (*nu=*)*py-r-ð* (<nous puerons>), mais la 3<sup>e</sup> personne du pluriel dans (*el=*)*py-r-ð* (<elles pueront>). Il s'agit alors d'exposants distincts, accidentellement homophones — ils ne l'étaient pas en ancien et moyen français avant la chute générale des consonnes finales. Il n'est pas pour autant nécessaire de les distinguer formellement — p.ex. /ð<sub>1</sub>/ et /ð<sub>2</sub>/ selon l'usage des lexicographes — puisqu'ils n'ont pas de réalité indépendante.

C'est aussi cette dépendance et que la relation forme-signification puisse n'être pas biunivoque — et l'est en fait rarement dans les langues dites « fusionnelles » (l'adjectif parle de lui-même) dont le latin et le français, FP comme FO, sont des spécimens — qui font que la distinction entre morphe et morphème perd son utilité, dès lors que les exposants, dégagés comme expliqué aux paragraphes précédents, s'associent directement à un nombre a priori indéfini de significations. Ce qui toutefois n'écarte pas la question du nombre d'exposants, qui est un problème empirique. Ainsi, à comparer le présent /pyð/ 'puons' et l'imparfait /pyjð/ 'puions' au niveau de leurs phonologies, on doit s'interroger quant au statut grammatical de /j/ : exposant signifiant « imparfait de l'indicatif » (d'où *py-j-ð*), ou partie d'un exposant /jð/ signifiant « 1<sup>ère</sup> personne du pluriel de l'imparfait de l'indicatif » (d'où *py-jð*) ? Je tâcherai de répondre au chapitre 9.

## 5.6. Les théories morphologiques

J'ai avoué au début de ce chapitre que l'hypothèse ici défendue de l'existence d'une composante morphologique autonome ne fait pas l'unanimité dans le petit monde des linguistes. De fait, personne ne nie l'existence de phénomènes morphologiques, ce serait nier l'évidence empirique. Le désaccord théorique porte sur le qualificatif « autonome ». Ce qui se passe à l'intérieur des mots — si tant est qu'il « s'y passe » quelque chose — obéit-il à des principes tout à fait ou en partie différents de ceux qui régissent les relations entre les mots, ou bien les mêmes opérations qui construisent des phrases sont-elles aussi à l'œuvre pour former des mots? La question, semble-t-il, ne s'est vraiment posée qu'à partir des années 50 du siècle écoulé. La raison en peut être que l'autre grande division de la grammaire, la syntaxe, la mise ensemble des mots, était jusqu'alors restée le parent pauvre des études

linguistiques (cf. Molino 1985 : 5-40). L'explication de ce retard gît peut-être en cela que la linguistique moderne est née au dix-neuvième siècle de l'étude des langues indo-européennes anciennes, sanskrit, latin, grec, gotique, etc., dont la morphologie flexionnelle est assez riche pour encoder la plupart des relations grammaticales, en sorte que l'agencement des mots et des syntagmes, la syntaxe donc, y jouit d'une grande liberté, au point d'y sembler inexistante en tant que système. Ou plutôt, si elle y est contrainte, c'est par des facteurs pragmatiques (mise en relief, pertinence pour l'énoncé, etc.) impossibles à décrire au moyen de règles comme celles qui ont fait le succès de la morphologie et de la phonologie.

Ceci n'étant pas une histoire de la linguistique, je me contenterai de rappeler le changement de paradigme (au sens de Kuhn) initié par Ferdinand de Saussure (1857-1913), qui fait passer le langage du statut d'organisme à celui de structure mentale (la parole) et sociale (la langue) ; l'influence certaine, aux Etats-Unis, de l'étude des langues amérindiennes par, entre autres, Franz Boas (1858-1942), Edward Sapir (1884-1939) et Leonard Bloomfield (1887-1949) ; enfin, en Europe et outre-atlantique, les travaux de Lucien Tesnière (1893-1954), Zellig Harris (1909-1992) et Noam Chomsky grâce auxquels la syntaxe devient structurale et mathématisable.<sup>77</sup> Il est remarquable que les études fondatrices de ces trois auteurs paraissent dans un créneau temporel étroit, 1959 (posthumément) pour le premier, 1951 pour le deuxième, 1957 pour le dernier.

Dès lors que la recherche sur les structures syntaxiques des langues naturelles se posait en tant que discipline scientifique — c'est-à-dire justiciable des critères qui font qu'un énoncé peut ou non être dit tel (voir par exemple Koyré 1973) — la question des rapports entre lesdites structures et les structures morphologiques — dont l'étude avait indiscutablement atteint un degré de rigueur qui la qualifiait comme science — ne pouvait que surgir. D'autant que, comme on l'a vu, l'interface entre les unes et les autres est significative.

A ce point, il convient de garder présente à l'esprit une distinction importante, je veux dire celle entre théorie et modèle. Une théorie est un corps d'hypothèses sur un sujet donné, qu'on s'efforce de démontrer ; un modèle est un ensemble d'algorithmes visant à produire cette démonstration dans le cadre d'une certaine théorie, revendiquée ou implicite. La confrontation aux « nouvelles syntaxes » (cf. Abeillé 1993 ; Rouveret 2015), la morphologie (flexionnelle) l'approchait divisée entre trois théories visant, au-delà de la description, à expliquer ce que « font » les sujets parlants quand ils fléchissent les mots de leurs langues (cf. Hockett 1954 ; Matthews 1972 ; Bybee 1985 ; Mel'čuk 1993, 1994, 1996 ; Spencer 1998 ; Fradin 2003 ; Haspelmath & Sims 2010). Elles sont connues sous leurs noms anglais *Item and Arrangement* (IA), *Item and Process* (IP) et *Word and Paradigm* (WP). Les « items » des deux premières désignent tout élément associant une forme quelconque (éventuellement imperceptible) et une signification. Dans la première, *Arrangement* 'arrangement' est à comprendre au sens de « disposition » « mise en place ». Ces théories divergent essentiellement sur deux points : (a) le statut de réalité des exposants ; (b) comment ils s'associent à ce qu'ils modifient.

Pour les tenants de l'IA, « les morphèmes sont des choses » (« *morphemes are things* ») (Spencer 1998 : xxx ; cf. aussi Fradin 1993). Donnant à « morphème » son extension maximale dans l'ordre linéaire, cela signifie (a) que bases et affixes sont à égalité des items lexicaux ; (b) qu'ils ne s'associent que par concaténation. *Quid* des procédés non-linéaires tels les pluriels *de facto* apophoniques (yiddish *fus / fis*) ou toujours métaphoniques (allemand *Fuss / Füsse*) ? Là, l'IA se conforme au principe selon lequel à chaque signification doit correspondre un élément (cf. figure x) et fait l'hypothèse de la présence dans l'un et l'autre cas d'un suffixe porteur de la pluralité, zéro en yiddish (*fis*+∅), audible en allemand (*fys*+ə). S'ensuit la nécessité de lister dans le lexique tous les allomorphes d'une base donnée.

---

<sup>77</sup> Cette liste est trop courte. Il faudrait, au minimum, évoquer la contribution de Richard Montague (1930-1971), mathématicien, philosophe et logicien américain (cf. Partee 2006).

A priori, la théorie ne requiert pas un procédé de concaténation particulier. C'est l'essor de la grammaire générative transformationnelle (GGT) qui fit germer l'idée de transposer à la flexion (et à la dérivation) les modes d'assemblage des constituants syntaxiques : structure hiérarchique arborescente, transformations par déplacement de constituants, têtes et projections fonctionnelles auxquelles attacher les éléments fonctionnels. On estimait par là simplifier la grammaire, externe aussi bien (par hypothèse) qu'intériorisée, puisqu'un seul type de mise en forme suffisait à expliquer toutes les expressions. Le premier modèle à brancher ainsi l'IA à la GGT fut sans doute la morphologie dite « distribuée » (*Distributed Morphology*) (cf. Halle & Marantz 1993), dont la version la plus récente est la nanosyntaxe (*Nanosyntax*) (cf. Starke 2009 ; De Clercq 2009 ; Collins & Kayne 2020). Ainsi, pour celle-ci, une expression finnoise telle que *karhuille* /karhu-i-lle/ {ours-pluriel-allatif} 'vers les ours' s'analyse comme un mini-arbre syntaxique (d'où « nano ») consistant en trois têtes fonctionnelles, K (cas), Num (nombre) et N (nom ou plutôt racine), et leurs projections KP, NumP et NP : [<sub>KP</sub> [<sub>K</sub> allatif (-lle)]] [<sub>NumP</sub> [<sub>Num</sub> pluriel (-i)]] [<sub>NP</sub> [<sub>N</sub> karhu-]]] (De Clercq 2009). Les morphèmes *-lle*, *-i-* et *karhu-* sont les items lexicaux qui « épellent » (*spell out*) les propriétés [allatif], [pluriel] et le signifié/concept *Ursus*. L'ordre dans lequel ils apparaissent s'obtient par application successive de la transformation « déplacer », qui adjoint NP à la gauche de NumP, puis la suite N-Num à la gauche de KP. On notera que la nanosyntaxe accepte les exposants cumulatifs (*portmanteau* 'valise') comme le latin *-as* de *puellas* 'jeunes filles' ou le *-um* de *canum*, pour lesquels les propriétés de genre (féminin ou masculin), de nombre (pluriel) et de cas (accusatif ou génitif) et les têtes fonctionnelles qui les portent sont dites « submorphémiques » (De Clercq 2009). Le principe « une signification, un élément » n'en est pas remis en cause, car les trois têtes fonctionnelles demeurent présentes dans la représentation, seul l'épel propre leur fait défaut.

L'IP a en commun avec l'IA d'accorder aux éléments (*items*) flexionnels une réalité indépendante, en sorte que les mots complexes fléchis peuvent résulter d'une simple concaténation, et en résultent en effet dans le cas de formations agglutinantes telles que l'exemple finnois donné à l'instant. Les deux théories divergent lorsque la simple concaténation ne suffit plus à décrire adéquatement la formation. Comme le nom l'indique, la notion de processus (*process*) est centrale en IP. Ainsi, plutôt que de lister deux bases pour l'allemand *Fuss* / *Füsse*, on dira — de façon, du reste, plus conforme à la tradition — que la suffixation de l'exposant de pluralité *-e* déclenche (ou est corrélatif à) un processus morphophonologique de relèvement de la voyelle du radical : /u/ → /y/. En yiddish, le processus analogue (/u/ → /i/) est seulement corrélé à l'opération sémantique de pluralisation, sans besoin d'un suffixe zéro. On suppose du même coup l'existence de bases ayant une réalisation par défaut (/fus/), mais susceptibles d'altérations phonologiques liées à telles modifications morphosyntaxiques et/ou sémantiques marquées. N'adhérant pas au principe « une signification, un élément », l'IP n'a aucun problème avec la flexion non-linéaire ni avec les exposants cumulatifs.

La troisième théorie, WP, est celle dont les racines plongent au plus profond dans l'histoire — elle remonte aux grammairiens antiques —, en même temps qu'elle sous-tend tout l'exposé des sections précédentes. Ses deux piliers sont (a) que les exposants flexionnels n'existent pas à part des mots (*words*) qui les contiennent, lesquels ne sont donc pas formés par concaténation, mais sont enregistrés tels quels dans le lexique mental ; (b) les mots fléchis entrent dans des paradigmes où ils signifient par contraste avec les autres entrées dont ils partagent tout ou partie de la forme (sauf supplétisme) et une partie de la signification. Par exemple, *šāt-š* 'chantons' apparaît formellement identique à *šāt-jš* 'chantions' jusqu'à un certain point (le phonème /t/), il dénote la même activité et vaut du même type d'agent (*nu* 'nous'), mais la temporalité est autre (le passé au lieu du présent) ; ou encore *māz-š* 'mangeons' : autre forme que *šāt-š* 'chantons' jusqu'à un certain point (le phonème /z/) et

autre activité dénotée, mais même temporalité et même type d'agent comme indiqué par le partage de la terminaison /ɔ̃/ dans le contexte d'une forme verbale. Ce sont ces contrastes entre formes à la fois identiques et distinctes qui amènent à isoler des parties stables ou variables et à insérer entre elles des frontières internes, d'où *fāt-ɔ̃* 'chantons' et *māz-ɔ̃* 'mangeons' (cf. Bybee 1988). Mais il s'agit là en tout état de cause d'une opération théorique, et rien n'autorise à affirmer a priori que parties et frontières trouvent un équivalent dans la connaissance naturelle (implicite, non réfléchie) que les sujets parlants ont des formes. Dès lors, la parcimonie alias rasoir d'Ockham impose de supposer jusqu'à preuve du contraire qu'au niveau de la connaissance et de l'usage de la langue les formes fléchies signifient de façon globale plutôt que par addition de significations partielles. Il n'est du reste pas difficile de trouver des cas où cette addition est de toute façon impossible : dans la périphrase temporelle  $z=\varepsilon$  *fāt-e* 'j'ai chanté', *avwar* a depuis longtemps coupé tout lien sémantique avec le verbe voulant dire « posséder » ; en fait, il ne veut rien dire, seule la périphrase entière a un sens.

Il s'ensuit que WP n'autorise aucun modèle du type de la nanosyntaxe. La théorie intègre le principe dit « d'intégrité lexicale » selon lequel les mots-formes sont des atomes (au sens antique et logique) par rapport aux contextes syntaxiques où ils s'insèrent, si bien qu'aucune opération syntaxique ne saurait intervenir à l'intérieur des mots et que morphologie et syntaxe constituent en effet deux composantes séparées et autonomes au sein de la grammaire, descriptive comme, par hypothèse, mentale (cf. Bresnan & Mchombo 1995).<sup>78</sup>

Ce caractère atomique du mot-forme n'interdit évidemment pas de l'analyser au niveau morphologique. Cela est d'autant plus vrai que, même si ses sous-parties n'ont pas d'existence autonome, les sujets parlants ont forcément conscience (ou subconscience) de la récurrence des mêmes objets phonologiques, les exposants, pour reprendre le terme introduit plus haut, exprimant les mêmes significations dans les mêmes contextes, par exemple, en espagnol, que /mos/ en finale d'un mot-forme verbal veut dire que le procès dénoté est le fait de plusieurs individus incluant la locutrice ou le locuteur : cf. *cantamos* '(nous) chantons', *cantaremos* '(nous) chanterons', etc. C'est bien ce qui rend les sujets parlants capables de conjuguer tout verbe encore jamais rencontré parce que néologique ou inventé.<sup>79</sup>

L'opposition la plus polaire passe donc entre IA et WP. Pour reprendre les termes de Fradin (2003), la première est morphématique — *morpheme-based* selon Haspelmath & Sims (2010) — tandis que la seconde est lexématique — *word-based* selon les mêmes auteurs.<sup>80</sup> La position d'IP apparaît médiane. L'opposition morphème / lexème ne suffit toutefois pas à classer l'ensemble des modélisations fondées sur l'une ou l'autre des trois théories.

Stump (2001 : 1-3) propose un classement sur la base de quatre qualités logiquement indépendantes : « lexical » (*lexical*), « inférentiel » (*inferential*), « incrémental » (*incremental*) et « réalisationnel » (*realizational*). Est lexical un modèle qui suppose que les exposants flexionnels sont inclus dans le lexique. (On voit à l'œuvre l'ambiguïté dénoncée à la note 60.) Est inférentiel un modèle pour lequel la présence de tel exposant, par exemple /s/ en anglais, résulte d'une règle associant à la base la signification dudit exposant, par exemple « 3<sup>e</sup> personne du singulier du présent de l'indicatif », d'où l'implication : {si verbe & 3SG.présent.indicatif  $\Rightarrow$  V-s}. Pour un modèle incrémental, la signification flexionnelle du mot — par exemple, que *likes* 'aime' est la 3<sup>e</sup> personne du singulier du présent de l'indicatif du verbe *like* — est le produit de l'ajout de l'exposant /s/ au radical, exposant qui peut aussi

<sup>78</sup> Il s'ensuit aussi que WP se marie au mieux avec les théories syntaxiques « surfacistes » (cf. Abeillé 1993 ; Abeillé & Godard 2021) qui n'ont pas l'usage des adjonctions et déplacements de la GGT et du minimalisme (Chomsky 1995).

<sup>79</sup> Les choses sont plus compliquées en FP. D'une part, on l'a vu, l'interprétation de /ɔ̃/ final dépend du temps : cf. *i(l)=fātərɔ̃* 'ils chanteront'. D'autre part, hors des verbes dits « réguliers » (1<sup>er</sup> et 2<sup>e</sup> groupes), les hésitations ne sont pas rares. (voir les chapitres 9 et suivants). Heureusement, les néologismes sont tous réguliers..

<sup>80</sup> On dit aussi « lexicaliste ». Mais le terme est source de malentendus, comme on va le voir.



bien être extrait du lexique (lexical) qu'impliqué par une règle (inférentiel). Un modèle réalisationnel procède à l'inverse : c'est la signification globale du mot fléchi qui légitime (*licenses*) l'apparition de l'exposant, que celui-ci jouisse ou non d'une individualité lexicale. Indépendantes, ces quatre qualités peuvent se combiner librement, ce qu'elles font en effet selon Stump. Pour se limiter à deux exemples, le modèle nanosyntaxique est clairement lexical, puisque les morphèmes y sont des items lexicaux (voir ci-dessus) ; il est en même temps réalisationnel, puisque les morphèmes concrets ne sont insérés qu'après que la (nano)syntaxe a assemblé les propriétés morphosyntaxiques idoines et les a adjointes à la base. On aura compris, en revanche, que tout modèle fondé sur WP sera inférentiel et réalisationnel.

Maintenant, si l'on considère comme la plus clivante la décision d'attribuer ou non une existence lexicale aux exposants flexionnels, on voit qu'un modèle non-lexical au sens de Stump (lexématique au sens de Fradin) ne peut qu'être inférentiel et réalisationnel, puisque les exposants flexionnels n'y ont pour abri que les mots-formes. D'où une nouvelle dichotomie entre modèles constructivistes (ou, mieux dit sans doute, assemblistes) et modèles abstractivistes. Pour les premiers, les mots fléchis sont des assemblages d'éléments qui leur préexistent nécessairement ; pour les seconds, comme expliqué plus haut, bases et exposants sont abstraits des mots-formes grâce aux relations paradigmatiques entre ceux-ci (cf. Blevins 2006).

Il est toutefois assez probable que cette dichotomie n'en soit pas vraiment une, dans la mesure où les deux possibilités ne s'excluent pas entièrement (cf. Stump 2019). Tout système flexionnel permet des généralisations, explicites si formulées par les linguistes, implicites si pratiquées par les sujets parlants. Celles-ci peuvent être dites abstractives si, comme dans l'exemple espagnol cité plus haut, forme et sens entretiennent une relation constante au sein d'une classe de mots donnée (*Vmos* = 1PL) ; ou bien s'il est possible de prédire telle forme d'un paradigme à partir de telle autre, par exemple, en FP, que le participe d'un verbe dont l'infinitif se termine par /e/ (*ŷate* 'chanter') se termine de même (*ŷate* 'chanté.e(s)'), ou, en allemand, que les verbes dont la désinence du prétérit inclut /tə/ (*ich glaubte* 'je croyais', *du glaubtest* 'tu croyais', etc.) ont un participe terminé par /t/ (*geglaubt* 'cru.e(s)'). Les généralisations constructives concernent la forme même des mots-formes, par exemple, toujours en allemand, que les verbes dont l'infinitif est en *-ieren* ne prennent pas le préfixe *ge-* au participe : *reformieren* 'réformer', *reformiert* 'réformé.e(s)' ; ou encore qu'en sanskrit, le locatif pluriel et le génitif pluriel demandent des radicaux différents : *rāja-su* 'chez les rois', *rājñ-ām* 'des rois' (Stump 2019) ; ou encore qu'en latin (sauf exception), si un mot a le même nombre de syllabes au nominatif et au génitif singuliers, alors son génitif pluriel présente la désinence *-ium*.

Au fond, comme je l'ai déjà laissé entendre, ces généralisations sont ce qui impartit aux exposants une visibilité et une certaine autonomie, sans qu'il soit en aucune façon nécessaire d'en faire des objets ayant une existence en dehors des lieux où ils se manifestent. Cela semble assez assuré s'agissant des exposants flexionnels. Les exposants dérivationnels, en tout cas certains d'entre eux, jouissent de davantage de liberté. Les préfixes tels que *macro-*, *micro-*, *super-*, etc. sont aussi des adjectifs : « au niveau macro », « c'est super » ; la terminaison *-isme* de *constructivisme*, *socialisme*, *macronisme*, etc. connaît un emploi nominal : *les ismes*, plus ou moins synonyme de « les doctrines ». Je n'imagine pas entendre des anglophones dire *the s's* pour *the plural nouns*. D'où une autre distinction qu'on doit faire, touchant celle-là au rapport des sujets parlants à leur savoir morphologique (cf. Kilani-Schoch & Dressler 2005). Il y a d'une part la connaissance des formes acquises dès l'enfance et fixées une fois pour toutes, par exemple que l'imparfait de *ʒə=bwa* 'je bois' est *ʒə=byve* 'je buvais', que *kaftjer* 'cafetière' est un mot simple nonobstant l'étymologie, etc. Cette connaissance ne s'accompagne probablement et, en tout cas, ne requiert aucune analyse

spontanée de la part de qui l’acquiert. Elle constitue ce qu’on peut appeler la morphologie statique. D’autre part, il y a la morphologie dynamique au moyen de laquelle les sujets parlants, s’appuyant sur leurs généralisations implicites, se montrent capables d’accueillir, voire de forger de nouvelles formes et de les intégrer aux schèmes (*patterns*) déjà connus. La distinction statique / dynamique n’est évidemment pas sans parenté avec l’opposition improductif / productif. Mais elle ne s’y confond pas, car relevant d’un autre niveau. Du point de vue du sujet parlant, le paradigme flexionnel de *bwar* ‘boire’ appartient à la morphologie statique. Reste que la flexion, par exemple l’association de la terminaison /ε/ à l’imparfait, est très généralement productive. De même, on peut parier que *bufəri* ‘boucherie’ est acquis comme un mot simple. Pourtant, la dérivation en /-əri/ est productive comme en témoigne le récent *krwasātəri* ‘croissanterie’. Il convient donc de ne pas confondre les formes elles-mêmes et leurs aléas avec les schèmes qu’il est possible d’en extraire par ces généralisations analogiques auxquelles on revient toujours.

On conclura de tout cela qu’un modèle WP équipé d’exposants flexionnels dotés du degré d’autonomie que leur confèrent les possibles généralisations abstraitives et constructives, explicites et implicites, ne diffère guère d’un modèle IP non-lexical. C’est un modèle de cette sorte qui sous-tendra les développements des chapitres suivants. Mon but étant de décrire, non de défendre une théorie — dont beaucoup de celles et ceux qui me liront n’ont, de toute façon et légitimement, que faire —, je ne chercherai pas à démontrer qu’aucun autre modèle ne parviendrait mieux à rendre compte des faits, voire n’en serait tout simplement capable. J’ai confiance que celles et ceux que la question intéresse s’en apercevront.

## 6. La flexion nominale

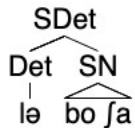
« Qu’est-ce que tu pratiques, comme sports ? demanda Bourrelrier en pensant le pluriel d’un s qu’il ne prononçait pas » (Queneau, *Les dimanches de la vie*, p. 175)

Dans leur majorité, les noms sont invariables en FP en cela qu’aucun affixe flexionnel ne s’y attache. La flexion procède par voie de cliticisation, elle est semi-analytique (voir chapitre 5). Une importante minorité connaît une flexion supplétive. Les propriétés flexionnelles nominales sont la définitude, la deixis, la possession, le genre grammatical et le nombre.

### 6.1. La définitude

On a vu au chapitre précédent pourquoi la définitude peut être considérée comme une propriété flexionnelle. Les articles défini et indéfini du FP sont des mots fonctionnels clitiques, si bien que *lə=fa* ‘le chat’ et *ẽ=fa* ‘un chat’ appartiennent au paradigme flexionnel du lexème CHAT, et cela même si, du fait de sa propriété syntagmatique, l’article s’appuie sur le premier mot du SN, quel qu’il soit, plutôt que sur le mot même qui réalise ledit lexème, pour lequel la propriété de définitude est pertinente et qu’on est convenu d’appeler la « tête » du syntagme (cf. chapitre 5) : cf. *lə=bo\_fa* ‘le beau chat’. Du point de vue morphosyntaxique, l’article défini se lie à l’adjectif prénominal, tandis que, du point de vue sémantique, c’est bien la tête *fa* ‘chat’ et, par conséquent, tout le SN [*bo\_fa* <beau chat>] qui est défini. (On note au passage que, loin d’être compact, ce type de SN est dislocable moyennant la prosodie adéquate : cf. *lə=bo, si ð=pø dir, fa* ‘le beau, si on peut dire, chat’.)

Il suit de ce caractère syntagmatique qu’il est parfaitement licite, voire nécessaire d’associer une structure syntaxique de constituants (c-structure) aux expressions définies ou indéfinies telles que *lə=bo\_fa* ‘le beau chat’ :



Le déterminant est la tête d'un syntagme déterminé SDet (en terminologie anglaise *Determiner Phrase*, DP) qui domine le SN. Mais cela ne contrevient en rien à la nature flexionnelle de la modification, c'est seulement que celle-ci, au lieu de procéder par voie synthétique (affixale) comme en arabe classique et en islandais, procède par voie semi-analytique ou, plus précisément, périphrastique : *lə=fa* 'le chat' est une périphrase nominale, de même que *a\_by (lə=ɛ)* 'a bu (le lait)' est une périphrase verbale (voir chapitre 9). (Voir aussi périphrastique *ply=move* 'plus mauvais' vs. synthétique *pir* 'pire'.)

La définitude est une propriété booléenne, c'est-à-dire limitée à deux valeurs, soit [+] (défini) et [-] (indéfini). La valeur générique (où le référent correspond à la totalité d'une classe naturelle) parasite quant à l'expression celle de la définitude positive ou négative : cf. *lə=fa ɛ(t) ẽ=mamifer* 'Le chat est un mammifère', *le=fa s̄\_de=mamifer* 'Les chats sont des mammifères', *ẽ=fa (nə=)mãz=pa\_də=fruɣi* 'Un chat (ne) mange pas de fruits' (mais voir plus loin). Les définitions de la définitude sont nombreuses et diverses. Je m'en tiendrai à une définition discursive. Sont affectés d'une valeur de définitude positive les SN — plus exactement les signifiés exprimés par eux — dont le locuteur considère (plus ou moins sciemment) que les référents en doivent être connus de son auditoire, car ou bien ils ont déjà été introduits dans l'échange en cours et sont ainsi « liés au discours » (*D-linked*), ou bien le sujet de l'échange doit les rendre évidents (cf. *lə=fa a\_ãkər\_fe\_yn=betiz* 'Le chat a encore fait une bêtise'), ou bien ils relèvent de ce que tout humain neurotypique et raisonnablement informé dans une situation donnée est censé capable d'identifier (la lune, le gouvernement, le covid, etc.). En français (FP comme FO) les SN présentant ces caractéristiques sont explicitement marqués comme positivement définis. Ceux qui ne les présentent pas sont explicitement marqués du contraire. Ne pas marquer un SN pour la définitude, positive ou négative, est possible, mais dans des constructions particulières : locutions telles que *prãdr\_gard* 'prendre garde', *done\_rez̄* 'donner raison', etc. ; énumérations : *vo, vaf, kof̄, kuve* 'Veaux, vaches, cochons, couvées' ; énoncés gnomiques (proverbes, adages, etc.) : *fjẽ ki=abwa nə=m̄r=pa* 'Chien qui aboie ne mord pas'. Ce sont les seuls cas où la généricité a sa marque propre, négative, par défaut : tout SN non marqué explicitement pour la définitude s'interprète par défaut comme référant à quoi que ce soit qui répond à la définition du lexème : tout veau, tout chien, etc. (voir 6.5 pour la pluralité, marquée en FO, non marquée en FP). Hors les contextes précis qui autorisent ce défaut, le marquage quasi constant de la valeur de définitude des SN est une particularité du français (FP et FO), que bien d'autres langues ne partagent pas. Sans même parler des langues, fort nombreuses, dépourvues de déterminants (articles) définis et indéfinis (japonais, russe, swahili, etc.), il suffit de parcourir un texte anglais pour être frappé de la quantité de SN « nus » là où le français ne pourrait se passer de déterminant. Traduit en anglais, ce paragraphe commence par « *Definiteness is a Boolean property...* ».

La définitude positive est exprimée en FP par un mot fonctionnel, un déterminant proclitique syntagmatique (article) attaché au premier constituant du SN, et dont une représentation lexématique concevable est L. Il cumule la propriété de définitude avec celles de genre et de nombre (cf. 6.4 et 6.5). Sa forme varie en fonction de ces propriétés, ainsi que de la nature consonantique ou vocalique de l'initiale de ce à quoi il s'attache.

Si la tête du SN est singulière et masculine et si le premier constituant du SN (qui peut en être ou non la tête) commence par une consonne ou un « *h* aspiré », l'article défini a pour

forme  $l\partial=$  :  $l\partial=zoli\_fa$  ‘le joli chat’  $l\partial=h\acute{e}tr$  ‘le hêtre’.<sup>81</sup> Avec une tête singulière et féminine et devant consonne ou *h*, la forme est  $la=$  :  $la=zoli\_fat$  ‘la jolie chatte’,  $la=haf$  ‘la hache’. Devant voyelle ou approximante à valeur vocalique et quel que soit le genre de la tête singulière du SN, on a  $l=$  :  $l=ot(r\acute{a})\_fa(t)$  ‘l’autre chat.te’,  $l=\acute{e}tr$  ‘l’être’,  $l=wazo$  ‘l’oiseau’,  $l=wa$  ‘l’oie’.<sup>82</sup> Qu’il s’agit bien d’allomorphie conditionnée, le démontre *a contrario* le fait que  $l\partial$  peut se réaliser /l/ devant consonne par amuïssement du schwa si le contexte phonologique s’y prête : p.ex.  $d\ddot{a}=l\partial=k\ddot{a}bwi$  ‘dans le cambouis’ /d\ddot{a}lk\ddot{a}bwi/, mais cela de manière facultative : /d\ddot{a}l\ddot{a}k\ddot{a}bwi/ est une prononciation possible. En revanche, \*/l\ddot{a}ot\ddot{f}a/ est tout à fait exclu, tandis que  $la=$  n’est jamais réalisé /l/ devant consonne. Au pluriel, l’article a pour forme  $le=$  et il est épïcène :  $le=zoli\_fa(t)$  ‘les joli.e(s) chat.te(s)’,  $le=h\acute{e}tr$  ‘les hêtres’,  $le=haf$  ‘les haches’. Devant voyelle ou approximante vocalique, un /z/ apparaît —  $le=z=ot(r\acute{a})\_fa(t)$  ‘les autres chat.te.s’,  $le=z=\acute{e}tr$  ‘les êtres’,  $le=z=wazo$  ‘les oiseaux’ — dont la nature sera étudiée en 6.5.

Devant consonne et *h*, les articles définis masculin singulier  $l\partial=$  et pluriel épïcène  $le=$  s’amalgament aux deux prépositions *a* ‘à’ et *d\partial* ‘de’ :  $o=p\acute{a}pl$  ‘au(x) peuple(s)’,  $o=travaj\acute{o}z$  ‘aux travailleuses’,  $o=h\acute{e}tr$  ‘au(x) hêtre(s)’,  $dy=p\acute{a}pl$  ‘du peuple’,  $dy=h\acute{e}tr$  ‘du hêtre’,  $de=travaj\acute{o}z$  ‘des travailleuses’,  $de=h\acute{e}tr$  ‘des hêtres’. Devant voyelle et approximante vocalique,  $l=$  masculin ou féminin ne s’amalgame pas :  $a=l=ot\acute{a}r$  ‘à l’auteur’,  $a=l=otris$  ‘à l’autrice’,  $d\partial=l=ot\acute{a}r$  ‘de l’auteur’,  $d\partial=l=otris$  ‘de l’autrice’,  $a=l=wazo$  ‘à l’oiseau’,  $d\partial=l=wazo$  ‘de l’oiseau’,  $a=l=wa$  ‘à l’oie’,  $d\partial=l=wa$  ‘de l’oie’. Attachée à un nom masculin, la forme *o* est ambiguë quant au nombre,<sup>83</sup> tandis que devant un féminin, elle ne peut être que plurielle, puisque l’article féminin singulier  $la=$  ~  $l=$  ne s’amalgame pas :  $a=la=travaj\acute{o}z$  ‘à la travailleuse’,  $d\partial=la=travaj\acute{o}z$  ‘de la travailleuse’ (voir ci-dessus pour  $l=$  féminin). Au pluriel, le même /z/ que plus haut apparaît devant voyelle :  $o=z=ot\acute{a}r$  ‘aux auteurs’,  $de=z=ot\acute{a}r$  ‘des auteurs’,  $o=z=otris$  ‘aux autrices’,  $de=z=otris$  ‘des autrices’ (cf. 6.5).

Aucune règle phonologique synchronique (raisonnable) ne permettant de produire /o/, /dy/ et /de/ à partir des suites plus transparentes mais agrammaticales \*/al\ddot{a}l/, \*/al\ddot{e}l/, \*/d\ddot{a}l\ddot{a}l/, \*/d\ddot{e}l\ddot{e}l/, je me risque à définir ces formes insécables comme des flexions de l’article défini. Je m’appuie pour cela sur l’analyse largement acceptée des prépositions *a* et *d\partial* comme des marqueurs casuels, *a* de datif, *d\partial* de génitif ou d’ablatif (*v\acute{a}nir d\partial=fe=swa* ‘venir de chez soi’),<sup>84</sup> d’où il ressort que, par exemple, *o* dans  $o=p\acute{a}pl$  exprime l’ensemble de propriétés {DEF +, GENRE *m*, NOMBRE [ ], CAS *datif*, VAL  $SN_{\#c}$ }, soit défini, masculin, sous-spécifié ([ ]) pour le nombre (c.-à-d. singulier ou pluriel), datif et sélectionnant (VAL pour valence) un SN dont la marge gauche consiste en une consonne (y compris *h*). Et l’on ne s’étonnera pas qu’une marque de flexion soit sensible pour sa forme à la phonologie de ce à quoi elle s’attache : en russe, par exemple, la désinence de génitif pluriel des masculins est *-ov* si le mot se finit par une consonne dure (p.ex. *vol-\acute{o}v* ‘des bœufs’, de *vol*), mais *-ej* s’il se finit par une consonne palatalisée (p.ex. *car-\acute{e}j* ‘des tsars’, de *car*). Une autre analyse, qui a la préférence de la GGF, consiste à définir *o*, *dy* et *de* comme des prépositions amalgamées. Elle a, me semble-t-il, l’inconvénient de mettre en avant la composante prépositionnelle de la forme, alors que la détermination me paraît contribuer au moins autant à la signification globale.

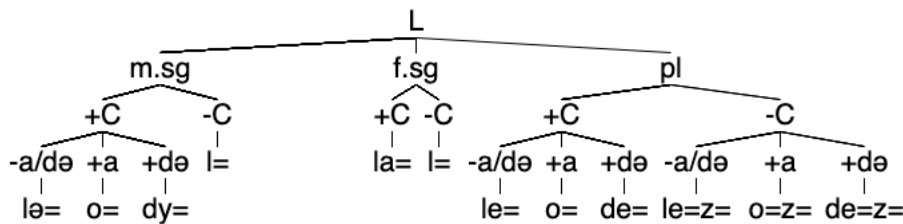
<sup>81</sup> Je représente le « h aspiré » par un *h*, qui n’est rien d’autre que la notation du fait que devant certains noms à initiale vocalique le schwa de l’article ne s’élide pas et il ne se fait pas de liaison. L’orthographe ne le note pas toujours : cf. *le onze* et non \**l’onze* ; *les onze* /le\ddot{z}z/, non \*/lez\ddot{z}z/.

<sup>82</sup> « Approximante » désigne ce qu’on appelle traditionnellement (et indistinctement) « semi-voyelle » ou « semi-consonne ». Comme on sait, les approximantes [j] et [w] ont valeur de consonne dans des mots encore sentis comme des emprunts : cf.  $la=jurt$  ‘la yourte’,  $l\partial=wistiti$  ‘le ouistiti’. On hésite pour certains :  $l=wat$  ‘l’ouate’ ou  $la=wat$  ‘la ouate’,  $l=j\acute{e}n$  ‘l’hyène’ (normatif) ou  $la=j\acute{e}n$  ‘la hyène’ (stigmatisé, mais commun).

<sup>83</sup> Voir au chapitre 1 la solution gaullienne à cette ambiguïté.

<sup>84</sup> « Cas » est pris ici au sens large qu’il a dans la Grammaire des Cas (*Case Grammar*, Fillmore 1968), à peu près synonyme de rôle sémantique.

L'arbre d'héritage ci-dessous résume les diverses formes de l'article défini du FP et les conditions de leur apparition :<sup>85</sup>



Bien entendu, seules les formes du pied de l'arbre sont dotées de réalité psychologique. Tout francophone voulant énoncer un état de fait exprimé par un terme de type nominal, positivement défini, caractérisé par telles propriétés de genre et de nombre et dans tels environnements phonologiques, sait que telle forme s'impose et l'emploie sans y penser. Comment elle ou il le sait, autrement dit comment ces informations sont encodées dans les réseaux neuronaux effectivement activés, voilà ce qu'on ne sait pas (voir chapitre 14).

Peut-être faut-il voir un autre amalgame dans *syl* qui fond l'article masculin singulier et la préposition *syr* 'sur'. Un argument en faveur de cette hypothèse est que, me semble-t-il, la restitution du schwa (p.ex. en discours ralenti par souci de clarté) entraîne forcément la réalisation du *r* : /syɾlətwa/, non \*/sylətwa/, mais /syɾltwa/ ou /sylvta/, d'où deux structures distinctes : *syr*=lə=twa et *syl*=twa 'sur le toit'. Mais il se peut aussi que la préposition *syr* ait *sy* pour allomorphe : cf. *sy*=la=fɛz ~ *syr*=la=fɛz 'sur la chaise'.

L'article UN exprime l'indéfinitude. Ses quatre formes sont supplétives : *ɛ*= dans un SN masculin singulier et devant consonne ou *h* : *ɛ*=zoli\_fa 'un joli chat', *ɛ*=hetr 'un hêtre' (IDF.M.SG) ; *ɛ̃n*= devant voyelle ou approximante vocalique : *ɛ̃n*=ot(rə)\_fa 'un autre chat', *ɛ̃n*=wazo 'un oiseau' ; *yn*= dans un SN féminin singulier, quelle que soit l'initiale : *yn*=zoli\_fat 'une jolie chatte', *yn*=ot(rə)\_fat 'une autre chatte', *yn*(ə)=haf 'une hache', *yn*=wa 'une oie', *yn*=jurt 'une yourte' (IDF.F.SG). Dans un SN pluriel de l'un ou l'autre genre, l'article indéfini est *de*= ou *də*= devant consonne et selon qu'un adjectif précède ou non le nom-tête : cf. *de* ~ *də*=zoli\_fa(t) 'de(s) joli.e.s chat.te.s' (IDF.M/F.PL), mais seulement *de*=fa(t) 'des chat.te.s' vs. \**də*=fa(t) — à moins qu'il ne s'agisse de l'article partitif (voir ci-dessous) ; *de*=z= devant voyelle ou approximante vocalique : *de*=z=an 'des ânes', *de*=z=wa 'des oies', alternant avec *d*= si un adjectif à initiale vocalique précède le nom-tête : cf. *d*=adorabl\_fa(t) ~ *de*=z=adorabl\_fa(t) 'd'adorables chat.te.s' ~ 'des adorables chat.te.s' (cf. 6.5 pour /z/). L'article indéfini féminin singulier *yn* est sans allomorphes.

Il existe enfin un article partitif D toujours accompagné de l'article défini dans la langue moderne et ayant pour formes *dy*= devant SN masculin singulier et initiale consonantique (*dy*=bδ\_taba 'du bon tabac'), *də*=la= devant SN féminin singulier et initiale consonantique (*də*=la\_bɔn\_sup 'de la bonne soupe'), *də*=l= devant SN masculin ou féminin singulier et initiale vocalique (*də*=l=er 'de l'air', *də*=l=o 'de l'eau').<sup>86</sup> Il s'associe à des SN dénotant des entités massives ou massifiées, c.-à-d. naturellement sans parties individualisables ou bien délibérément conçues comme telles (cf. *də*=la=bel\_tomat 'de la belle tomate') et de ce fait exemptées de l'opposition de nombre. C'est sans doute ce qui explique que, devant SN pluriel (*de*=tomat 'des tomates'), *de* 'des' représente plutôt l'article

<sup>85</sup> Lire « -a/də » comme « partout ailleurs qu'après *a* 'à' et *də* 'de'.

<sup>86</sup> Le français classique avait 'de' pour 'du' devant un adjectif prénominal, comme dans la chanson bien connue — mais est-ce encore vrai ? — « J'ai de bon tabac dans ma tabatière ». Cette tournure est désormais archaïque. Plutôt obsolète ou, à tout le moins littéraire, me semble-t-il, est l'emploi de l'article partitif avec l'article indéfini dans des énoncés généralement exclamatifs : *il*=ε *d*=yn=betiz ! 'Il est d'une bêtise !' Mais cela reste à vérifier : *il*=ε *d*=ɛ̃=kδ ! 'Il est d'un con !' me paraît assez banal.

défini pluriel qu'une forme de l'article partitif.<sup>87</sup> D'où trois points de vue sur la réalité, dont la différence est aussi évidente qu'elle est difficile à expliciter en peu de mots : *lə=bḥ taba* (*ε kã mem movε pur=la=sãte*) 'Le bon tabac (est quand même mauvais pour la santé)' vs. (*ʒ=ε lṣtã fyme*) *ḗ=bḥ taba* 'J'ai longtemps fumé) un bon tabac' vs. (*fyme*) *dy=bḥ taba* (*n=evit=pa lə=kãser dy=pumḥ*) '(Fumer) du bon tabac (n'évite pas le cancer du poumon)'. La forme *d(ə)* s'emploie en contexte négatif : cf. *pa\_d=o* 'pas d'eau', *ply\_də=sigar* 'plus de cigares' — mais, dans *plys\_də=sigar* 'plus (davantage) de cigares', a-t-on affaire à l'article partitif ou à la préposition *də* 'de' ? Même question pour *mwē\_də* 'moins de', *boku\_də* 'beaucoup de', *otã\_də* 'autant de' etc., tous items qui se rangent dans la catégorie des quantificateurs, dont la généralisation inclut les déterminants (cf. Barwise & Cooper 1981 ; Keenan & Stavi 1986). La distinction entre article indéfini pluriel, article partitif et préposition *də*, toutes formes en partie homophones, est donc une question complexe, mais qui ne nous concerne pas directement (cf. Kupferman 1979, 1994).

Cette section est aussi l'endroit où mentionner le déterminant interrogatif-exclamatif *kəl* 'quel.le(s)' invariable en genre et nombre si le successeur commence par une consonne : *Kəl=fa eskə ty=prefer* ? 'Quel(s) chat(s) est-ce que tu préfères ?', *Kəl=fās t(y)=a* ! 'Quelle chance tu as !'. S'il est vocalique et pluriel, apparaît le même /z/ qu'après les autres déterminants : cf. *Kəl=z=anəri* ! 'Quelles âneries !', *Kəl=z=wazo* ? 'Quels oiseaux ?' (cf. 6.5 ; voir le chapitre 7 pour l'adjectif homophone).

## 6.2. La deixis

Au sens étroit où je la prends ici, la deixis a pour domaine le déterminant déictique (démonstratif) proclitique CE, qui se réalise sous trois formes : *sə* 'ce', *set* 'cet' ou 'cette', et *se* 'ces'. Le premier est masculin singulier devant consonne et *h* : *sə=ʒoli\_fa* 'ce joli chat', *sə=hetr* 'ce hêtre' (DEM.M.SG) ; *set* est féminin singulier devant consonne : *set=ʒoli\_fat* 'cette jolie chatte' (DEM.F.SG) ou bien épïcène devant voyelle ou approximante vocalique : *set=adorabl\_fa(t)* 'cet.te adorable chat.te', *set=wazo* 'cet oiseau', *set=wa* 'cette oie' (DEM.M/F.SG) ; *se* est épïcène et pluriel : *se=ʒoli\_fa(t)* 'ces joli.e.s chat.te.s' (DEM.M/F.PL) — et toujours le /z/ devant voyelle ou approximante vocalique : *se=z=adorabl\_fa(t)* 'ces adorables chat.te.s', *se=z=wazo* 'ces oiseaux', *se=z=wa* 'ces oies' (DEM.M/F.PL). Devant voyelle, *set=* se réalise communément /st/ : /stadɔʁabləʃa(t)/. La prononciation /stœ/ devant consonne (/stœʃat/) relève d'un registre « bas ». La distance physique ou discursive peut se marquer au moyen des particules enclitiques syntagmatiques *si* et *la* : *sə=fa\_nwar=si* 'ce chat noir-ci' (proche) vs. *sə=fa\_nwar=la* 'ce chat noir-là' (distant).

## 6.3. La possession

En tant que propriété flexionnelle, elle est marquée par une série de déterminants possessifs proclitiques semblables aux articles et déictiques examinés ci-dessus. Ces déterminants varient de forme selon la personne du possesseur (pr), le genre grammatical et le nombre de l'entité possédée ou possessum (pm) et l'initiale consonantique ou vocalique du mot qui les suit : avec pr = 1|2|3SG, pm = M.SG, on a *mḥ|tḥ|sḥ=kanari* 'mon|ton|son canari', mais *mḥn|tḥn|sḥn=an* 'mon|ton|son âne', *mḥn|tḥn|sḥn=wazo* 'mon|ton|son oiseau'. L'examen de la phonologie historique du français peut laisser penser que l'alternance [ḥ] vs. [ḥn] relève d'un processus phonologique plutôt que du supplétisme — solution qu'on ne saurait toutefois écarter en synchronie. Avec pr = 1|2|3SG, pm = F.SG, on a *ma|ta|sa=pi* 'ma|ta|sa pie', mais *mḥn|tḥn|sḥn=alwet* 'mon|ton|son alouette', *mḥn|tḥn|sḥn=wa* 'mon|ton|son oie', où l'on voit que ces déterminants gardent la forme masculine devant un possessum féminin dès lors que le mot qui l'exprime commence par une voyelle ou une approximante vocalique. Il s'agit bien alors de supplétisme par rapport à la forme féminine préconsonantique.

<sup>87</sup> De même devant les noms toujours pluriels (*pluralia tantum*) : *de=sizo* 'des ciseaux'. On peut hésiter avec *de=z=epinar* 'des épinards' (exemple de Grevisse 2009).

Le tableau 1 présente l'ensemble des paradigmes. Dans les colonnes 1|2|3SG et 1|2PL, je ne donne que le paradigme de la 1<sup>ère</sup> personne. Les paradigmes 2|3SG (*tǝ, sǝ*) sont identiques à 1SG moyennant la différence de consonne initiale ; de même le paradigme 2PL (*vǝt(r(ǝ))*, aussi forme de politesse) par rapport à 1PL.

possesseur				
possessu m		1 2  3SG	1 2PL	3PL
	M.SGC	mǝ=	nǝt(r(ǝ))=	lǝr=
	M.SG V	mǝn=		
	F.SGC	ma=		
	F.SGV	mǝn=		
	PLC	me=	no=	
	PLV	me=z=	no=z=	lǝr=z=

Tableau 1 : les déterminants possessifs en FV (C = devant consonne, V = devant voyelle ou approximante vocalique).

Il est tentant — d'aucuns ont cédé — d'analyser les formes du singulier : la consonne initiale exprimerait la personne, la voyelle le genre et le nombre. Cette analyse ne se justifie pas, me semble-t-il. Il est vrai que /m/, /t/ et /s/ se retrouvent dans les pronoms personnels correspondants (voir le chapitre 8), et /a/ et /e/ dans les articles définis *la* 'la' et *le* 'les'. Mais /ǝ(n)/ est isolé. En fait, ces rapprochements sont le produit de l'histoire : les formes phonologiques /mǝ(n)/, /ma/ et /me/, par exemple, remontent aux formes latines *meum*, *meam* et *meos*, respectivement accusatif masculin singulier, accusatif féminin singulier et accusatif masculin pluriel d'une dérivation adjectivale à partir du radical /me/ du pronom personnel de 1<sup>ère</sup> personne. Les formes latines sont analysables ; les françaises ne le sont plus. Ce qui n'interdit pas de supposer que les résonances qu'on y perçoit en puissent faciliter l'acquisition et la mémorisation.

Question déjà posée à propos de l'article pluriel : quel est le statut grammatical du /z/ que manifestent toutes les formes dont l'hôte est membre d'un SN pluriel et commence par une voyelle ou une approximante vocalique ? Liée qu'elle est à la question plus générale du marquage de la pluralité en FP, je tâcherai d'y répondre sous 6.5.

Le rapprochement des déterminants définis, déictiques et possessifs révèle une constante : le caractère épïcène de la forme plurielle.

Les pronoms possessifs *mǝ* | *mǝn* 'mien.ne(s)', *tǝ* | *tǝn* 'tien.ne(s)', *sǝ* | *sǝn* 'sien.ne(s)', *nǝt* 'nôte(s)', *vǝt* 'vôte(s)', *lǝr* 'leur(s)' ne relèvent pas de la flexion, même dans leur usage comme épithètes, que le FP contemporain ne pratique que par jeu — p.ex. *ǝ=mǝ kuzǝ* 'un mien cousin' — et de façon très figée quant aux possessums acceptables : *ǝ=mǝ ǝa* 'un mien chat' me paraît à la limite du compréhensible. La présence obligatoire du déterminant indéfini ou déictique (*sǝ=mǝ kuzǝ* 'ce mien cousin') démontre qu'on a bien affaire à des adjectifs, non à des déterminants. (Pour une raison que j'ignore, *lǝ=mǝ kuzǝ* 'le mien cousin', bien formé en français classique, ne se dit plus. S'écrit-il encore ?) Quant à *sǝ=livr ǝ mǝ* 'Ce livre est mien', il est décidément archaïque, il ne s'emploie « que par badinage » dit Grevisse (1980, §906). Dans *sǝ=livr ǝ lǝ=mǝ* 'Ce livre est le mien', banal en revanche, mais qui présuppose un contexte de choix entre plusieurs candidats à la possession, l'adjectif est substantivé.

#### 6.4. Le genre grammatical

Il présente deux valeurs, masculin et féminin, en principe associées au genre biologique, mais de fait arbitraires pour les noms dénotant des entités non sexuées : *la=tabl* 'la table', *lǝ=kabl*

'le cable',  $\tilde{\epsilon}$ =*virys* 'un virus', *yn=bakteri* 'une bactérie' ; ou bien sexuées, mais trop éloignées de la faune prototypique (sous nos latitudes) pour que la distinction soit jugée pertinente : *la=ziraf* 'la girafe', *lə=kāguru* 'le kangourou', mâle ou femelle. Sans parler des animaux dont le sexe est difficilement reconnaissable : *lə=kərbo* 'le corbeau', *la=sigal* 'la cigale', etc.

Seul un sous-ensemble des noms dénotant des êtres animés possède une flexion de genre. Il en existe deux sortes. D'une part, des noms comportant une consonne finale « latente » et alternant de ce fait une forme courte et une forme longue affectée au féminin. La flexion de genre de ces noms est donc identique à celle des adjectifs variables traitée en détail au chapitre 7. Je ne donne ici que quelques exemples :  $\tilde{\epsilon}n=avoka$  'un avocat' / *yn=avokat* 'une avocate',  $\tilde{\epsilon}n=ekrivē$  'un écrivain' / *yn=ekriven* 'une écrivaine',  $\tilde{\epsilon}=polisje$  'un policier' / *yn=polisjer* 'une policière',  $\tilde{\epsilon}=fa$  'un chat' / *yn=fat* 'une chatte', etc. D'autre part, il y a les paires telles que *naʒær* 'nageur' / *naʒøz* 'nageuse', *direktær* 'directeur' / *direktris* 'directrice', *an* 'âne' / *anes* 'ânesse', dont la formation est traditionnellement considérée comme étant affaire de morphologie dérivationnelle plutôt que flexionnelle. Il semble toutefois plus exact d'admettre, à la suite de Bonami & Boyé (2019), que *direktær* 'directeur' / *direktris* 'directrice' et les autres paires de ce type sont bien deux formes d'un même lexème en relation flexionnelle, mais des formes parallèles en ce sens qu'aucune des deux n'est dérivée de l'autre. Il s'ensuit, résultat intéressant, que les deux sortes de noms à genre manifeste (*gender iconic*) participent du même système, avec la même relative imprévisibilité de la relation formelle entre masculin et féminin. On ne peut donc que supposer que les enfants les acquièrent de même en parallèle, sans possibilité d'abstraire une règle, comme le font probablement les petit.e.s Allemand.e.s face à une langue où, au contraire, les noms de personnes féminins sont régulièrement dérivés des masculins : cf. *Kanzler* / *Kanzlerin* 'chancelier / chancelière', *Lehrer* / *Lehrerin* 'instituteur / institutrice'. Quant aux noms de profession comme *profesær* 'professeur', l'usage se répand de leur donner un féminin au moyen d'un <e> orthographique (« une professeure »). Je les considère comme invariables en FP (*yn=profesær*), tout autant que les quelques adjectifs en -ær tels que *mejær* 'meilleur.e', *mazær* 'majeur.e', etc. pour lesquels cet <e> est de tradition.

Rien dans la forme des noms simples n'en indique le genre, mis à part une certaine corrélation statistique entre telles finales (non significatives) et l'un ou l'autre genre grammatical. Ainsi, les mots terminés en /ð/ sont majoritairement masculins — mais *la=mezð* 'la maison' et bien d'autres exceptions (cf. Corbett 1991). Pour les noms dérivés dont je ne traite pas, en revanche, le genre grammatical est systématiquement associé à telles terminaisons : les noms en .*asjð* sont féminins (*la=rymin.asjð* 'la rumination'), ceux en .*aʒ* sont masculins (*lə=lav.aʒ* 'le lavage'), etc. Avec les noms à deux genres manifestes (*avoka* 'avocat' / *avokat* 'avocate', *naʒær* 'nageur' / *naʒøz* 'nageuse', *direktær* 'directeur' / *direktris* 'directrice'), la finale est censée indiquer le genre. Mais elle n'y réussit qu'associée à l'information que le nom dénote un individu sexué et possède une autre forme, partiellement distincte et désignant un individu du même type, mais de l'autre sexe : *patat* 'patate' est féminin, mais sa finale /at/ n'y entre pour rien ; une *pðsøz* 'ponceuse' n'est pas l'équivalent féminin d'un ?*pðsær* 'ponceur' — la profession existe-t-elle seulement ? La coexistence stable, imposée aux locuteurs (plutôt que produite à volonté par eux) de deux formes est cruciale. Sans elle une Bouriate pourrait être la femme d'un inexistant (mais formable) \*Bouriat, comme une bougnate peut être celle d'un bougnat, couple disparu du paysage, mais toujours lexicalement vivant. Cela étant, il va de soi qu'ayant assimilé un certain nombre de paires lexicalisées, les acquérants ne manqueront de s'essayer à des créations analogiques qui rencontreront ou non l'assentiment de l'entourage. S'il y a des nageurs et des nageuses, il y a des chances pour qu'il y ait aussi des influenceurs et des influenceuses ; si des directeurs et des directrices, des instituteurs et des institutrices. Mais pourquoi pas des \*/diʁektøz/ ou, plus simplement, des \*/diʁiʒøz/, ou encore, plus improbable, des \*/diʁektøʁes/ sur le modèle des



quelques féminins archaïques du type d' *ã/ãtərəs* 'enchanteresse' ? C'est là que l'usage fait sentir son poids, car on peut douter que la règle qui permet de discriminer les féminins en /øz/ des féminins en /tɔis/ — existe-t-il ou non un verbe de même radical ? — soit accessible aux enfants acquérant la langue.

Ce tableau complexe a pour conséquence qu'en tant que trait flexionnel explicitement marqué de l'une ou l'autre de ses valeurs, le genre grammatical n'est souvent porté ou à tout le moins clairement manifesté en FP que par les modifieurs accordés aux noms, les déterminants (voir plus haut) et certains adjectifs (voir le chapitre 7).

### 6.5. Le nombre

L'attribut de nombre a deux valeurs : singulier et pluriel. Par défaut, sauf indice du contraire et de fait dans la majorité des cas, cette valeur n'est indiquée que par la forme du déterminant qui accompagne presque toujours le nom : p.ex. singulier. *lə=fa* 'le chat', *sə=fa* 'ce chat', *ẽ=fa* 'un chat', *la=fat* 'la chatte', *set=fat* 'cette chatte', *yn=fat* 'une chatte' ; pluriel *le=fa(t)* 'les chat.te.s', *se=fa(t)* 'ces chat.te.s' *de=fa(t)* 'des chat.te.s'. Le morphe graphique <s> du FO est sans équivalent en FP — et l'on peut douter que Bourrelier, tout en sachant qu'il voulait parler de plusieurs sports, l'ait « pensé ». <sup>88</sup> Sauf que, comme indiqué plus haut, un *z* suit les déterminants pluriels devant voyelle ou approximante vocalique : *le=z=an* 'les ânes', *le=z=wa* 'les oies', *de=z=an* 'des ânes', *de=z=wa* 'des oies', *me=z=an* 'mes ânes', *me=z=wa* 'mes oies', *no=z=an* 'nos ânes', *no=z=wa* 'nos oies', *lær=z=an* 'leurs ânes', *lær=z=wa* 'leurs oies'. <sup>89</sup> Et de même les adjectifs prénominaux des SN pluriels dans les mêmes conditions : *me=vjø=z=ami* 'mes vieux amis', *me=vjej=z=wa* 'mes vieilles oies'. Ce *z* récurrent a-t-il un statut grammatical et, si oui, lequel ?

Le fait est que sa présence ne paraît pas nécessaire pour marquer la pluralité dans presque tous les exemples ci-dessus. Mais elle l'est pour le contraste entre *lær=an* 'leur âne' et *lær=z=an* 'leurs ânes' où le possessif ne varie ni en genre ni en nombre, si bien que *z* apparaît comme le seul indice de l'opposition de nombre. Et aussi pour *o=zæn=ami* 'au jeune ami' opposé à *o=zæn=z=ami* 'aux jeunes amis'. Il ne peut donc s'agir d'une consonne « de liaison » ne servant qu'à éviter un hiatus indésirable pour quelque raison, comme dans *don=mwa•z=ã*, variante réputée plus « familière », stigmatisée, mais courante de *don=m=ã* 'donne-m'en', ou comme il s'en trouve dans d'autres langues d'oïl, p.ex. en poitevin-saintongeais *Doune me z-ou* /dunmøzu/ 'Donne-le moi' (Arantéle). <sup>90</sup> D'autant que l'hiatus se tolère fort bien dans /leaʃ/ ('les haches') et que la question ne se pose pas dans *lær=z=an* 'leurs ânes' et *zæn=z=ami* 'jeunes amis'.

Il n'est donc pas douteux que *z* soit associé au trait [NOMBRE *pluriel*]. C'est un morphe, c'est-à-dire l'association d'une forme phonologique et d'une signification grammaticale, mais doté de propriétés particulières. Au sein d'un SN pluriel il est susceptible de s'attacher devant tout constituant à initiale vocalique : *lær=z=ãsjê\_z=ami* /lœkzãsjêzami/. Comme on voit, en présence de plusieurs constituants à initiale vocalique, *z* peut s'attacher à chacun d'eux. Cette attache est à peu près invariable quand l'élément le précédant est un déterminant ; elle est changeante quand il s'agit d'un adjectif antéposé (voir plus loin). Si aucun constituant ne remplit cette condition, le SN peut être ambigu : cf. *lær=bel\_vwatyr* 'leur(s) belle(s) voiture(s)', qui ne dit pas si les gens dont il est question possèdent une ou plusieurs voitures.

Mais faut-il vraiment s'étonner de la présence de ce *z* ? L'exposant du trait [NOMBRE *pluriel*] en français comme dans les autres langues romanes occidentales n'est-il pas un

<sup>88</sup> Il est vrai que les personnes immergées dans l'écrit, dont l'auteur de cet ouvrage, ont tendance à voir mentalement les mots qu'elles prononcent en même temps qu'elles les prononcent. Elles sont probablement minoritaires. Et ce ne peut être le cas des enfants non scolarisés.

<sup>89</sup> Quand /j/ et /w/ ont valeur consonantique, *z* n'apparaît pas : cf. *de=jurt* 'des yourtes', *de=wistiti* 'des ouistiti'.

<sup>90</sup> Dans *Pās•z=i* 'Penses-y', le FO s'est vue obligée d'intégrer la consonne de liaison vu l'impossibilité de \*/pãsi/.

suffixe *-s*, voisé en position intervocalique ? Pourquoi le /z/ entendu dans /lezan/ ('les ânes') et /lœrzan/ ('leurs ânes') serait-il autre chose que ce suffixe, amuï par les accidents de l'histoire, mais dont les contextes de liaison révèlent encore la présence ? C'est là l'explication communément reçue, on l'a vu au chapitre 3. Elle ne va pas sans problèmes.

Le premier, évident et que l'explication traditionnelle élude, est que, hors de ces contextes, ce suffixe n'a plus de réalité qu'à l'écrit. Alors qu'en ancien français (dixième-quinzième siècles) tous les /s/ finaux se faisaient entendre, ils ont cessé en partie de se prononcer au cours du seizième siècle, en même temps que l'ensemble des consonnes finales. Au départ, cet amuïssement partiel est phonologiquement conditionné. Citons Gougenheim (1951 : 31) : « La prononciation des consonnes finales est régie au XVI<sup>e</sup> siècle par une loi de phonétique syntaxique : les consonnes finales se prononcent devant un mot commençant par une voyelle (c'est notre « liaison ») et à la pause (c'est-à-dire devant un arrêt du discours) ; elles ne se prononcent pas devant une consonne ». Une phrase telle que *Nous sommes tous tenus de prier Dieu* se prononçait donc /nusɔmtutəny dəprijedjø/, où ne se faisait entendre que le dernier /s/ du premier mot phonologique (Henri Estienne 1528?-1598, cité par Gougenheim 1951 : 32). On comparera avec le FP contemporain /nusɔmtustœny.../, où le *s* de *tus* 'tous' s'est lexicalisé et *tœny* 'tenu.e(s)' est invariable. (Les numéraux *sis* 'six' et *dis* 'dix' ont conservé cette règle, du moins en prononciation normative : cf. *Nu=sɔm sis* /nusɔmsis/ 'Nous sommes six', *si=vwatyr* /sivwatyr/ 'six voitures', *siz=ami* /sizami/ 'six amis'. Mais le /s/ tend à la fixité : on entend *sis=vwatyr* /sisvwatyr/.)

Il est par ailleurs remarquable que, dans le domaine gallo-roman, cette évolution n'est pas propre aux langues d'oïl, mais touche aussi les variétés nord-occitane et provençale de l'occitan, avec la même conséquence que le pluriel n'y est marqué à l'oral que sur les déterminants : cf. nord-occitan (limousin) *lo buòu* /lu'bjaw/ 'le bœuf' vs. *los buòus* /loj~low~ly'bjaw/ 'les boeufs' (Bec 1973 : 153).<sup>91</sup> Le nord-occitan n'en a pas moins préservé un marquage du nombre sur le nom, mais qui ne met pas /s/ en jeu : si le nom se termine par un /e/ atone, celui-ci peut se diphtonguer : cf. *lo ase* /lu'aze/ 'l'âne' vs. *los ases* /low'azej/ 'les ânes' ; le /o/ final des féminins singuliers se change en /a/ au pluriel : cf. *la filha* /lɔ'fijo/ 'la fille' vs. *las filhas* /la'fija/ 'les filles'. En occitan provençal, comme en oïl, l'article défini pluriel est épïcène et le nom ne se fléchit pas phonétiquement : cf. *lo flume* /lu'flyme/ 'le fleuve', *la vaca* /la'vako/ 'la vache' vs. *li flumes* /li'flyme/ 'les fleuve', *li vacas* /li'vako/ 'les vaches' (Bec 1973 : 118). Cet /i/ de l'article pluriel se retrouve à la finale des adjectifs féminins prénominaux : cf. *una polida femna* /ynopu'lido'feno/ 'une belle femme' vs. *de polidi femnas* /depu'lidi'feno/ 'de belles femmes'. Phénomène remarquable, un /z/ s'insère quand le nom commence par une voyelle : cf. *de pichòtis abelhas* /depi'tsotiza'bejo/ 'de petites abeilles'. Or, nonobstant l'orthographe, Bec (1973 : 116) précise que « ... cet -s est un pur phonème de liaison et n'a rien à voir diachroniquement (même s'il interfère avec elle) avec la marque du pluriel. » L'occitan provençal a donc suivi une évolution parallèle à celle du français, sans toutefois la pousser aussi loin, si bien qu'on pourrait opposer un occitan orthographique à un occitan parlé, n'était cette différence essentielle par rapport au français qu'il existe une variété importante, l'occitan languedocien et gascon, où l'écart entre graphie et phonologie reste minime et où le pluriel de *lo buòu* /lu'bjəw/ est *los buòus* /luz'bjəws/, de telle sorte qu'écrit et oral relèvent de la même typologie.

Tant qu'on a continué à le prononcer devant pause, le /s/ de pluriel a pu préserver sa réalité psychologique ; les enfants l'acquerraient en même temps que le reste de la flexion de la langue orale du seizième siècle. Mais les premiers signes du déclin se faisaient sentir. Gougenheim (1951: 33) note qu'aussi bien en contexte de liaison qu'à la pause, /s/ tombait souvent après un schwa final, toujours prononcé en moyen français : cf. *des prunes blanches et noires* /deprynəblāʃənwerə/ au lieu de /deprynəblāʃəzənwerəs/ ; et que les contemporains

<sup>91</sup> La graphie adoptée par l'auteur pour toutes les variétés est celle de l'Institut d'études occitanes (ieo-occ.org).

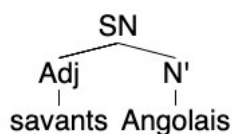
attribuaient cet usage aux « dames de Paris », bien qu'il fût probablement plus répandu.<sup>92</sup> Le changement était probablement accompli au dix-septième siècle, comme en témoigne cette phrase (combien juste !) que donne Brunot (1939: 216) citant la *Grammatica Gallica* de Kohlhans (1667) : « Les ignorants et les présomptueux sont altiers, mais les sachants modestes et humbles » /lezijorãzeleprezõtjyø sõtaltjɛ melesafã modɛst(œ)zɛœblœ/.<sup>93</sup> Les liaisons sont bien préservées, mais ni le -s (<x>) de pluriel de « présomptueux », ni celui de « humbles », tous deux devant pause, ne se prononcent.

La disparition de la pause comme facteur de réalisation et, sans doute, le rétrécissement des contextes de liaison ont ainsi eu pour effet de raréfier les apparitions de -s de pluriel, ce qui n'a pu qu'en modifier le statut dans la grammaire intériorisée des francophones. A quoi il faut ajouter que le voisement lui est devenu propre et ne peut plus s'expliquer par des causes phonologiques : les *s* finaux fixes (non pluriels) ne se voient pas entre voyelles : cf. *fis\_êdij* 'fils indigne' /fisêdij/, non \*/fizêdij/.<sup>94</sup>

Deuxièmement, la liaison en général et en *z* en particulier est un phénomène complexe en FP, pour lequel la traditionnelle tripartition prescriptive entre obligatoire, facultative et interdite ne vaut pas. L'étude de corpus de Mallet (2008) montre que seule la liaison entre déterminant pluriel et voyelle initiale est réalisée dans cent pour cent des cas ou peu s'en faut. Pour le reste, la variation est grande, en particulier s'agissant de la liaison à partir d'un adjectif antéposé ou vers un adjectif postposé. La prescription est qu'on doit lier dans *de savants Angolais* /dəsavãzãgolɛ/ (Adj < N), alors qu'on n'y est pas obligé dans *des savants angolais* /desavã(z)ãgolɛ/ (N < Adj). L'accord est général quant au caractère obligatoire du premier cas (de l'adjectif vers le nom). Les auteurs ne sont pas unanimes, en revanche, quant au second (du nom vers l'adjectif). Certains (p.ex. Delattre 1966) considèrent la liaison comme en effet facultative ; d'autres (Grammont 1914/1963 : 132 ; Milner & Regnault 1987) la rejettent et n'acceptent que /desavããgolɛ/ avec hiatus. (On verra plus loin pourquoi \*/desavãtãgolɛ/ est exclu.)

Ce désaccord est sans doute dû au fait que l'usage oral standard s'est modifié sur ce point vers la fin du dix-neuvième siècle. Même dans le registre le plus « châtié » du FP contemporain /dəsavãzãgolɛ/ ('de savants Angolais') paraît affecté — sauf à en faire le premier hémistiche d'un alexandrin classique. Le même désaccord, dû sans doute à la même cause, s'étend au cas où le nom pluriel est suivi d'un syntagme verbal dont il est en général le sujet : Delattre accepte *Ses plans ont réussi* /seplã(z)õkɛysi/, avec liaison variable, alors que Grammont et Milner & Regnault ne veulent que /seplãõkɛysi/ avec hiatus. A nouveau, la prononciation avec liaison sonne pédante, vieux jeu — y compris, me semble-t-il, en diction poétique.

Il se peut que la différence entre « avant le nom » et « après le nom » soit liée à une différence de structure syntaxique. Les éléments précédant le nom (déterminant, adjectif) appartiennent sans conteste possible au SN dont le nom est la tête :

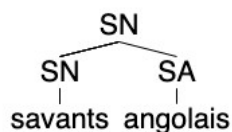


<sup>92</sup> Le rôle moteur des femmes dans la diffusion des innovations linguistiques est bien connu (cf. Labov 2001).

<sup>93</sup> Je convertis en API la transcription phonétique de Kohlhans reproduite par Brunot.

<sup>94</sup> Il n'en va pas ainsi en catalan où /s/ fixe se voise dans les mêmes conditions que *s* pluriel : *calç i guix* /kãlzi 'gij/ 'chaux et plâtre' comme *els arbres* /əl'zãbrəs/ 'les arbres' (Fabra 1941 : 23). Il n'en a pas toujours été ainsi en français non plus, car /fizêdij/ est la prononciation que recommande Littré, qui veut aussi qu'on prononce /fi/ dans *le fils de Dieu* (Milner & Regnault 1987 : 158). Le FP ne l'a pas suivi. Au pluriel (FO *des fils indignes*), une prononciation châtiée, voire hypercorrecte /defis(ə)zêdij/ s'entend parfois. Il est plus courant d'éviter la liaison : /defisêdij/.

Ce qui suit le nom relève d'un autre syntagme, à l'évidence s'il s'agit d'un SV comme dans l'exemple de Delattre, de façon plausible si l'on admet que l'adjectif postposé forme un syntagme adjectival (SA) adjoint au SN, comme schématisé ci-dessous :



L'équivalence généralement constatable entre syntagme et mot phonologique amène alors Milner & Regnault (1987 : 50) à décréter une « loi générale » selon laquelle « Il n'y a pas de liaison d'un mot phonologique au mot phonologique suivant ». On peut voir là une reformulation, pas forcément meilleure, du principe de Grammont (1914/1963 : 129) : « on lie dans l'intérieur d'un élément rythmique, on ne lie pas d'un élément rythmique au suivant ». C'est aussi la thèse défendue par Selkirk (1974) et critiquée par Morin & Kaye (1982).

Quoi qu'il en soit, tout cela n'a guère de poids en FP : on ne fait quasiment jamais entendre de /z/ en discours spontané du nom vers l'adjectif, et variablement de l'adjectif vers le nom, selon le degré de surveillance du discours et aussi, plus décisif, selon que le nom et l'adjectif antéposé forment ou non une unité sémantique d'interprétation non compositionnelle, ce qu'ils font dans bien des cas, vu les conditions qui président à l'antéposition des adjectifs (Nølle 1996) : cf. /lep(ə)ti(z)ami/ 'les petits amis' (de taille non précisée), /mezās̃jẽ(z)ami/ 'mes anciens amis' (qui n'en sont plus), /mevjø(z)ami/ 'mes vieux amis' (qui le sont toujours, mais ne sont pas forcément des vieillards), etc., tous cas de liaison très fréquente, quasi constante, mais néanmoins variable, comme le montre Mallet (2008). On comparera encore /legråd(z)ilyzjð/ 'les grandes illusions', variable, et /legrāzəm/ 'les grands hommes', lié dans près de cent pour cent des cas (cf. Mallet 2008).

Ajoutons encore que l'insertion de *z* paraît très fréquente, dès lors que la phonologie s'y prête, si elle permet seule de faire entendre la pluralité : cf. /oʒœnzami/ 'aux jeunes amis' vs. /oʒœnami/ 'au jeune ami'. (Mais l'ambiguïté de /oʒœnkopẽ/ 'au(x) jeune(s) copain(s)' est grammaticalement irrémédiable.) Plus variable, en revanche, quand la pluralité s'exprime par ailleurs : cf. mezœn(z)ami/ 'mes jeunes amis'.

La troisième raison de douter du caractère suffixal de *z* provient de ce que j'appellerai le « *z* migrateur » — puisqu'aussi bien notre hypothétique *-s* pluriel ne s'entend jamais que voisé — que la tradition nomme « velours », tel l'exemple suivant relevé par Morin & Kaye (1982 : 321) : 'des avions à réaction *z*-américains' (/dezavjðæeksjð\_zamevikẽ/) où *z* marque l'accord en nombre de l'adjectif avec le composé pluriel 'avions à réaction'.<sup>95</sup> Dans le registre du FP le plus soumis au contrôle du FO (FP standard), ces *z* fourvoyés sont sévèrement bannis et prononcer /z/ n'est admis — pour autant que ce le soit — qu'en situation de contiguïté : *des avions américains* réalisé /dezavjð(z)amevikẽ/.<sup>96</sup> Dans toutes les variétés, le fait est que *z* se manifeste le plus souvent quand le constituant qui le précède immédiatement partage le trait de pluralité avec celui qui le suit immédiatement (cf. /lœrzās̃jẽ(z)ami/ 'leurs anciens amis'). Mais rien n'interdit que cette concordance se fasse à distance, comme dans /dezavjð...zamevikẽ], où les constituants accordés sont séparés par le constituant non pluriel /æeksjð/. L'apparente rareté du phénomène, si les nécessaires études de corpus la confirment, est sans doute due au fait qu'aucun registre n'échappe tout à fait à la surveillance du FP standard. Mais ils s'entendent et ne sont pas si saugrenus. Il y a aussi le *z* qui se glisse entre les numéraux à syllabe finale fermée et les noms vocaliques : cf. /sẽkzāfā/ 'cinq *z*-enfants', /iksœzāfā/ 'X *z*-enfants' ; ou après certains quantifieurs : cf. /tʁodzāfā/ 'trop de *z*-enfants' ; et aussi entre préfixe et base comme dans /deminizordinatœr/ 'des mini-*z*-ordinateurs',

<sup>95</sup> La liaison au sein des composés (/avjðzæeksjð/) est une zone de variation que je laisse de côté.

<sup>96</sup> Anachronisme à part, Henri Estienne aurait prononcé /dezavjðzamerikẽs/. La première liaison s'est faite variable et la condition « pause » a disparu.

/desiberzatak/ ‘des cyber-z-attaques’, etc. (cf. Morin & Kaye 1982 : 323 ; observation personnelle pour le dernier exemple). La possibilité même de telles formes suffit à démontrer que *z* ne se rattache pas au constituant qui le précède à la façon d’un suffixe, équivalent FP du -s FO que le contexte intervocalique est censé faire remonter à la surface. Cette hypothèse, envisageable pour /avjõzamerikẽ/, est exclue pour /sẽkzãfã/ et cas analogues : *sẽk* ‘cinq’ se termine par une consonne. Qui plus est, le trait morphosémantique bivalent [pluriel] n’est pas approprié aux adjectifs cardinaux qui, sauf 1, mais y compris, paradoxalement’, 0 (cf. *zero\_zurno* ‘zéro journaux’), signifient par eux-mêmes une pluralité ; ni aux préfixes quasi autonomes *mini-* ‘mini-’ ou *siber-* ‘cyber-’ ; ni à un adverbe quantifieur comme *tro(da)*.

Mais si le morphe *z* n’est pas la manifestation orale du suffixe graphique pluriel du FO, qu’est-il ? Gougenheim (1938) y voit un préfixe attaché au nom dans les séquences déterminant-nom ou à l’adjectif dans les séquences nom-adjectif, où, d’accord avec Delattre, il accepte une liaison facultative. Aussi analyse-t-il *les amis* en /le z-ami/ et (*des*) *nez aquilins* en /ne z-akilẽ/ (en variation avec /ne akilẽ/). On notera que cet exemple confirme la réalité de *z* morphe associé au pluriel : le *z* orthographique de *nez* singulier ne se manifeste jamais : *un nez aquilin* /ẽneakilẽ/, non \*/ẽnezakilẽ/.<sup>97</sup> Morin & Kaye (1982), reprenant l’analyse de Gougenheim, attirent de même l’attention sur le contraste entre *un marchand de draps anglais* /...dʁazãglẽ/ où il paraît clair que ce sont les draps qui sont anglais, *un marchand de drap(s) anglais* /...dʁaãglẽ/ où c’est plutôt le marchand qui l’est — mais l’autre interprétation reste possible — et *des marchands de draps anglais* /...dʁazãglẽ/ à tout coup ambigu quant à savoir qui ou quoi est anglais, des marchands ou des draps — et les deux peuvent l’être.

Pour surprenante qu’elle puisse paraître, l’analyse préfixale est confirmée en partie par des études psycholinguistiques sur l’acquisition de la liaison par l’enfant (cf. Côté 2005). Elles montrent que les enfants de deux à trois ans encodent systématiquement les consonnes de liaison, quel qu’en soit le statut grammatical, en les rattachant à l’initiale du mot suivant, en sorte que leur lexique en formation contient pour un temps plusieurs variantes d’un même mot, par exemple /nuks/ et /züks/ pour *urs* ‘ours’, d’où des formes « déviantes » telles que /ẽzüks/ ou /denuks/. Mais les études indiquent aussi que, dans la suite du développement, ces consonnes se détachent et deviennent épenthétiques, c’est-à-dire insérées dans la séquence sans être rattachées lexicalement ni au mot précédent ni au suivant.

On s’étonnera peut-être de voir un morphe être ainsi sensible pour sa présence — car pour la forme le fait est banal — à la phonologie de son hôte, puisque *z* ne peut apparaître que devant un mot vocalique. Mais le phénomène est loin d’être propre au FP. Je n’en donnerai que trois exemples, mais il est à parier qu’on en trouverait d’autres. En tchéchéne, langue caucasienne du nord-est, les noms se répartissent entre plusieurs classes nominales (ou genres) exprimées par des préfixes consonantiques (*v-*, *b-*, *d-*, etc.) qui s’attachent, non pas au nom lui-même, mais au verbe dont le nom est le sujet ou l’objet selon que le verbe est intransitif ou transitif (Partchieva & Guérin 1996 : 71-72).<sup>98</sup> Mais cette attache n’est possible que si le dit verbe commence par une voyelle ; s’il commence par une consonne, l’accord ne peut s’exprimer (voir note 17).

Le fa d’Ambô, créole portugais parlé sur l’île d’Annobón dans le Golfe de Guinée, marque la pluralité nominale de deux façons : en préposant au nom le marqueur *nan* (apparenté au pronom 3PL) ; en préfixant *z-*.<sup>99</sup> Le premier procédé est réservé aux noms dénotant des humains et quelques animés non-humains intégrés à la société humaine (*nan pay* ‘pères, hommes’, *nan khasô* ‘chiens’) ; le second, lexicalement limité, est insensible à la distinction animé / inanimé, mais il n’est utilisable que si le nom commence par une voyelle

<sup>97</sup> Si on l’entend, ce que je me garderai d’exclure, c’est par hypercorrection.

<sup>98</sup> Le tchéchéne est une langue dite « ergative » (cf. Creissels 2006a : 314 et suiv.). Plus précisément, donc, le verbe s’accorde en classe avec son argument marqué du cas absolutif, c.-à-d. sujet d’un verbe intransitif ou objet d’un verbe transitif.

<sup>99</sup> Fa d’Ambô signifie ‘langue d’Annobón’. Annobón est une province de la Guinée-Equatoriale.

(*ome / zome* ‘homme(s), mari(s)’, *aba / zaba* ‘branche(s)’). « *This entails that inanimate nouns may never be marked for plural if they start with a consonant* » (Hagemeyer et al. 2020 : 35), si bien que le nombre ne peut s’exprimer que si le nom est accompagné d’un déterminant, d’un quantifieur ou d’un adjectif cardinal. Evidemment, ce préfixe *z-* provient du suffixe *-s* des articles définis portugais *os* (M.PL) et *as* (F.PL) ‘les’, voisé quand intervocalique : *os homens* [u'zomẽjʃ] ‘les hommes’ > *zome* ‘hommes’.

Mon troisième exemple est peut-être encore plus pertinent, car il s’agit d’un créole français, le réunionnais, qui présente lui aussi un /z/ initial associé au marquage du pluriel, qui n’est évidemment pas sans rapport avec le *z* du FP.<sup>100</sup> A la différence du *z-* du fa d’Ambô, il n’apparaît pas sur les noms à initiale vocalique, mais sur ceux qui, au singulier, commencent par un /l/ issu de l’article *l’* amalgamé, auquel il se substitue : p.ex. *in lékol* ‘une école’ vs. *zékol* ‘(des) écoles’, *in linfirmyé* ‘un.e infirmier/ère’ vs. *zinfirmyé* ‘(des) infirmiers/ères’ (Caron 2011 : 68-69). Les noms dont le /l/ ou /z/ initial appartient au radical de l’étymon français ne sont pas concernés : p.ex. *lima* ‘limace’, *ziz* ‘juge’ — non plus du reste que bon nombre de noms ayant intégré l’article français singulier ou pluriel à la base : p.ex. *loto* ‘auto, voiture’, *zabèy* ‘abeille’. Ces noms ou bien ne se pluralisent pas — *zabèy* peut faire référence à un nombre quelconque d’abeilles, de 1 à *n* — ou bien le font par préposition de la marque de pluriel *bann* /bã/ (< *bande*) : cf. *bann ziz* ‘les juges’, *bann loto* ‘les voitures’.

Pour Caron (2011), les noms comme *lékol* et *linfirmyé* comportent deux préfixes en alternance : *l-* exposant du trait [NOMBRE *sg*] et *z-* exposant du trait [NOMBRE *pl*]. Cette analyse se heurte toutefois au caractère très variable du phénomène. Il n’existe pas, semble-t-il, de locuteurs pour qui la forme en /l/ serait toujours singulière et celle en /z/ toujours plurielle. C’est affaire de préférence, plus ou moins affirmée selon les locuteurs et les lexèmes. On parlera donc plutôt d’allomorphie radicale, c.-à-d. d’alternance entre deux radicaux /lX/ et /zX/ facultativement associés aux deux nombres. L’appartenance d’un lexème donné au groupe susceptible de cette allomorphie est imprédictible.

On se gardera de conclure des deux derniers exemples que je considère le FP comme une sorte de créole. En aucune façon. Les conditions sociolinguistiques pour qu’il en soit un sont totalement absentes. Je ne les donne que pour montrer que les changements qui mènent au FP contemporain n’ont rien d’exceptionnel, sont même attendus vu les conditions initiales. On a vu que le FP est même allé plus loin que le fa d’Ambô sur la voie de l’autonomie de *z*.

Quant aux détails de ces changements, ils ne se laissent évidemment pas retracer avec précision. Le fait est qu’on ne trouve nulle trace de *z* migrateur dans les textes anciens, sans surprise, puisqu’il s’est toujours agi, semble-t-il, d’un phénomène exclusivement oral. Tout ce qu’on peut dire avec quelque certitude, c’est que le changement crucial dut être la raréfaction des liaisons dites « facultatives », jusqu’au point où la liaison avec *-s* de pluriel ne subsista pratiquement que là où son absence aurait entraîné une agrammaticalité (cf. Encrevé 1988; Bonami et al. 2005). Ainsi, dans les trois premiers mots de la sentence de Kohlhans citée plus haut, seule la première liaison (/lezijnorã/) est assurée ; il n’est pas exclu que la deuxième (/lezijnorãz/), quasi inexistante en FP contemporain, ait déjà été en voie d’extinction dans le parler ordinaire de la seconde moitié du dix-septième siècle. Maintenant, une particularité importante de *-s* (/z/) en tant que marque de pluriel sur les noms est d’être très généralement redondant, puisque la pluralité s’exprime aussi et distinctement par la forme des déterminants rarement absents. Cette redondance a pu contribuer à sa chute ; il se peut aussi qu’elle ait aidé à sa libération, assistée en cela par deux autres emplois du phonème /z/.

Il y a d’abord, naturellement, les liaisons dues aux /z/ finaux qui ne sont pas des marques de pluriel, obligatoires comme dans *døz=ami* ‘deux amis’, ou bien seulement possibles — et donc le plus souvent absentes en FP — comme dans *ƶ=ave(z)\_ê=rev* ‘J’avais un rêve’ (cf. Encrevé 1988). Quant aux liaisons à partir d’un numéral, il est d’ailleurs

<sup>100</sup> Sur l’origine du réunionnais, cf. Chaudenson (1974, 2003).

remarquable qu'un mot récemment entré dans la langue comme *oro* 'euro' se comporte pour beaucoup de francophones comme s'il commençait par *h* : cf. *di=oro* /diøʁo/ 'dix euros', alors qu'on n'entend pas, me semble-t-il, */?leøʁo/*. Il est vrai que la prononciation des numéraux est depuis longtemps une zone de variation particulièrement complexe, qui sort de mon domaine.

Mais surtout il y a la consonne antihiatu (paragogique) qu'on entend, en dépit de sa condamnation, dans *Dɔn=mwa=z=ã* ! 'Donne-moi-z'en !' pour *Dɔn=m=ã* ! 'Donne m'en !'. Il est du reste des cas où il paraît difficile de s'en passer : qui veut énoncer l'impératif correspondant à *Ty=l=i=ãmen* 'Tu l'y emmènes' n'a guère d'autre choix que *Ãmen=lə=z=i* ! 'Emmène-le-z'y !', car *\*/ãmenləi/* enfreint la contrainte qui empêche schwa d'apparaître devant une voyelle, tandis qu'*Ãmen=l=i* ! 'Emmène-l'y !', quoique acceptable, a quelque chose de surané (cf. Milner & Regnault 1987: 87).<sup>101</sup> Comme on sait, l'orthographe s'est vue contrainte d'enregistrer certains /z/ antihiatu, ainsi dans *Vas-y* !, dont le <s> est dépourvu de toute substance morphologique ou étymologique (cf. Frei 1929/2007). En outre /z/ antihiatu était plus commun, car peu voire pas du tout stigmatisé, au dix-septième siècle qu'il ne l'est de nos jours (Brunot 1939). Commun, il l'est aussi dans les autres langues d'oïl telles le poitevin: cf. *Donne me z-ou* ! /dunməzu/ 'Donne-le moi !' (Arantéle) ou *I sé pa s'i va o-z-aconsentir* /isepa siva ozakðsâtir/ 'Je ne sais pas si je vais l'accepter' (Gautier 1986: 29).

On ne peut donc exclure que le *z* pluriel migrateur constitue une extension du /z/ antihiatu, que les locuteurs n'auraient eu qu'à emplir d'un contenu sémantique, à savoir ajouter le trait [NOMBRE pl] au patron phonologique /XV#VY → XVzVY/, où V veut dire « voyelle », X et Y sont des suites phonologiques quelconques, et /z/ commence par combler l'hiatus (#), puis prend son sens si l'expression se trouve dans la portée d'un opérateur de pluralité. Ce processus, s'il est avéré, se serait produit au sein d'une dérive générale vers l'invariabilité morphologique des noms, elle-même liée à la multiplication des cas où le suffixe de pluriel se voyait privé de contrepartie phonologique du fait de la désuétude des liaisons non obligatoires.

J'adopte donc globalement l'analyse de Gougenheim et Morin & Kaye. Ma seule réserve est quant à la définition exacte de *z* : plutôt qu'un préfixe, j'y vois un proclitique, ce qu'indique ma notation. Cette définition me paraît plus adéquate compte tenu des données psycholinguistiques qui en révèlent, semble-t-il, le caractère épenthétique. Il est vrai, je l'ai dit, que la distinction entre affixe et clitique est souvent subtile, voire floue et que le critère le plus fiable est l'attache phonologique : les affixes s'attachent en principe au constituant auquel ils confèrent les traits qu'ils expriment. Les clitiques sont plus volages, ils sont susceptibles de s'attacher à des constituants avec lesquels ils n'entretiennent pas de relation sémantique. Ce n'est pas le cas de *z* : même en cas d'accord à distance comme dans *avjð.a.reaksjð\_z=amerikẽ* il s'attache bien à l'adjectif dont il démontre la pluralité ainsi que de tout le SN. D'une façon générale le français ne présente pas de données convaincantes permettant d'appliquer ce critère. Plus pertinent est le fait que les affixes sont généralement dédiés à telles parties du discours : nominaux (p.ex. les désinences casuelles du latin), verbaux (p.ex. *-e* de *fât-e* 'chantais / chantait'), etc. Or *z* apparaît devant des noms, mais aussi devant des formes verbales avec la même fonction d'exprimer la pluralité (cf. chapitre 9). La précision est importante, car elle montre qu'il ne s'agit pas d'homonymie comme pour le *-s* anglais, pluriel si suffixé à un nom, présent 3SG si à un verbe, génitif si à un SN. Je persiste donc à voir en *z* un proclitique plutôt qu'un préfixe, tout en reconnaissant volontiers que rien d'important n'en découle.

L'invariabilité interne des noms et des adjectifs quant au nombre — autrement dit, l'absence de flexion synthétique de nombre — représente le cas par défaut en FP. Défaut dont s'écarte un sous-ensemble assez important de noms et adjectifs masculins terminés en [al] à

<sup>101</sup> *Ãmen=z=i=lə* ! 'Emmène-z'y-le !', marginalement possible, exige aussi /z/.

pluriel en [o] : *ẽ=zurnal rezjonal* ‘un journal régional’ / *de=zurno rezjono* ‘des journaux régionaux’, auxquels s’ajoutent sept noms masculins terminés en [aj] : *baj* ‘bail’ / *bo* ‘baux’, *emaj* ‘émail’ / *emo* ‘émaux’, *koraj* ‘corail’ / *koro* ‘coraux’, *supiraj* ‘souponirail’ / *supiro* ‘souponiriaux’, *travaj* ‘travail’ / *travo* ‘travaux’, *vãtaj* ‘ventail’ / *vãto* ‘ventaux’, *vitraj* ‘vitrail’ / *vitro* ‘vitraux’.<sup>102</sup> Il n’y aurait pas grand sens à considérer [o] comme un morphe de pluriel dans *zurno* et ces sept noms, pas davantage que dans *boko* ‘bocaux’, *zenero* ‘généraux’, *fãvo* ‘chevaux’, etc., vu l’inanité grammaticale des reliquats /zurn/, /b/, /em/, /bok/, /zener/, /fãv/, etc.<sup>103</sup> Il s’agit donc là de supplétisme, dont la version totale est illustrée par *æj* ‘œil’ / *jø* ‘yeux’, tandis que les paires *bæf* ‘bœuf’ / *bø* ‘bœufs’, *æf* ‘œuf’ / *ø* ‘œufs’, et *ɔs* ‘os’ / *o* ‘os’, en font voir une version partielle — du reste en voie d’extinction, un nombre croissant de francophones n’énonçant jamais que *de=bæf*, *de=z=æf*, *de=z=ɔs*. Et qui dit encore *de=ser* pour ce qui s’écrit *des cerfs* ?

Les noms et adjectifs féminins terminés en [al] et [aj] sont invariables : *de=bakanal kolosal* ‘des bacchanales colossales’, *de=bataj zeneral* ‘des batailles générales’. De même les noms masculins en [aj] autres que les sept listés ci-dessus et de nombreux noms et adjectifs masculins en [al] tels que *fakal* ‘chacal’, *festival* ‘festival’, *regal* ‘régal’, *naval* ‘naval’, *tonal* ‘tonal’, etc.

En résumé, il n’existe pas en FP de flexion affixale de pluriel telle qu’en présentent les autres langues romanes. Dans leur majorité les noms et les adjectifs ne marquent pas la pluralité. Ceux qui la marquent le font par supplétisme partiel (*zurnal* / *zurno*, *bæf* / *bø*) ou total pour une unité (*æj* vs. *jø*). (Dans *æf* / *ø* et *ɔs* / *o*, le relèvement de la voyelle survivante en syllabe ouverte est automatique, /œ/ et /ɔ/ n’étant pas des mots possibles. On peut donc voir là aussi un supplétisme partiel.) Les déterminants marquent toujours la pluralité, également par supplétisme partiel (*lã/la* vs. *le*) ou total (*ẽ/yn* vs. *de*). Le morphe *z* est un proclitique dont l’insertion dépend de la forme phonologique du constituant qui le suit immédiatement (initiale vocalique ou semi-vocalique) et n’est quasi constante que s’il est précédé d’un déterminant pluriel.

Doit-on en conclure que, dans leur majorité, les noms du FP sont non spécifiés pour le nombre, ou, pour le dire autrement, que le nombre n’est pas une propriété appropriée pour la plupart, comme ce semble être le cas de tous dans des langues comme le chinois mandarin ou le japonais ? La réponse est oui et non. Non, car le FP ne diffère pas des autres langues romanes, FO compris, quant au fondement sémantique de l’attribution du nombre aux noms comptables — c.-à-d. aux noms dénotant des entités manifestées par des individus (ou exemplaires) distincts susceptibles d’être dénombrés : cf. « Il y a 1/2/3...*n* chaise(s) dans cette pièce ». Les noms comptables se distinguent ainsi des noms dits « massifs » dont les référents existent sous forme de masses indifférenciées : cf. « Il y a de l’eau/du sable, etc. dans cette pièce » (cf. Nicolas 2002). Au niveau cognitif-pragmatique, cette attribution est réglée par un principe qu’on peut qualifier d’arithmétique, en vertu duquel il est toujours nécessaire d’indiquer formellement si la cardinalité (le nombre d’exemplaires) du référent est égale ou supérieure à 1 : {*|n|* = 1 ⇒ singulier}, {*|n|* > 1 ⇒ pluriel}. Par là, le FP et ses congénères se distinguent des nombreuses langues où, dans un contexte donné, discursif ou textuel, cette indication peut n’être fournie que si on la suppose pertinente compte tenu du savoir partagé (ou tenu pour tel) par les interlocuteurs ou l’auteur/autrice et son lectorat. Par exemple, à la question « Quel métier fait Marie ? », on peut répondre en haïtien *Li vann kay*, littéralement « Elle vend maison », sans attacher à *kay* ‘maison’ aucune marque de nombre ou d’une quelconque autre propriété morphosyntaxique, confiant que tout un chacun comprendra

<sup>102</sup> *Fermaj* ‘fermail’ / *fermo* ‘fermaux’ et *zẽmaj* ‘gemmail’ / *zẽmo* ‘gemmaux’ sont des termes techniques peu connus. *Aj* ‘ail’ / *o* ‘aulx’ s’achemine vers l’archaïsme.

<sup>103</sup> /supir/ et /vittr/ correspondent bien à des lexèmes existants, *supir* ‘soupir’ et *vittr* ‘vitre’, mais *supiraj* et *vittraj* ne s’y rattachent sans doute pas en synchronie, vu l’éloignement des dénnotations.



qu'elle n'en vend pas qu'une seule. Dans ces langues, il est grammaticalement possible d'utiliser des mots-formes « nus », qui n'expriment rien de plus que leur signification lexicale. En français, FP comme FO, en revanche, cette possibilité est exclue, sauf dans quelques cas particuliers comme nous le verrons. Dans celui qui nous occupe, la grammaire n'autorise pas cette économie, on ne peut que répondre  $\mathcal{E}l=v\tilde{a} de=mez\tilde{\delta}$  'Elle vend des maisons'. Elle n'autorise pas non plus, à l'instar du portugais européen ou de l'anglais, les pluriels sans déterminants : *Vende casas*, *She sells houses*, le plus souvent non reconnaissables en FP ( $*\mathcal{E}l=v\tilde{a} mez\tilde{\delta}$ ,  $*\mathcal{E}l=v\tilde{a} zurno$ ), mais a priori toujours formables et lisibles en FO ( $*\langle$ Elle vend maisons $\rangle$ ,  $*\langle$ Elle vend journaux $\rangle$ ). Mais l'égale malformation de  $*Vende casa$  et  $*She sells house$  montre que ces deux langues se conforment comme le français au principe arithmétique.

On voit ainsi que ce qui sépare le FP du FO n'est pas le principe du marquage, inchangé, mais son procédé : en FP, le marquage repose entièrement sur le déterminant proclitique, ici l'article indéfini pluriel  $de=$  ; en FO, il se partage entre ce même article et le suffixe graphique  $\langle -s \rangle$ . Autrement dit, la flexion de nombre du FP est semi-analytique, celle du FO synthétique — je ne me prononce pas sur la nature morphosyntaxique de  $\langle des \rangle$ . Il s'ensuit que le paradigme d'un lexème comme MAISON dénotant une entité comptable comporte quatre cases en FP, nombre et définitude indissociables (voir plus haut) et le genre une propriété du lexème :

	singulier	pluriel
défini	$la=mez\tilde{\delta}$	$le=mez\tilde{\delta}$
indéfini	$yn=mez\tilde{\delta}$	$de=mez\tilde{\delta}$

Paradigme de FP  $mez\tilde{\delta}$

En revanche, si l'on considère que l'association des déterminants aux noms en FO est pure affaire de syntaxe — autrement dit que  $\langle des maisons \rangle$ ,  $\langle la maison \rangle$ , etc. ne sont pas des périphrases nominales telle que définies plus haut, mais uniquement des syntagmes [Det N] — le paradigme du même lexème ne comporte que deux cases en FO :

singulier	pluriel
$\langle maison \rangle$	$\langle maisons \rangle$

Paradigme de FO  $\langle maison \rangle$

Le premier paradigme vaut également des noms du type de *zurnal*, puisque le pluriel *zurno* ne peut s'analyser que comme une forme supplétive, associée au pluriel, mais ne l'exprimant pas par elle-même. En FO, en revanche, le radical supplétif  $\langle journaux \rangle$  est conditionné par le morphe pluriel  $\langle x \rangle$ , si bien que  $\langle journal \rangle$  /  $\langle journaux \rangle$  relève du même paradigme que  $\langle maison \rangle$  /  $\langle maisons \rangle$ , moyennant l'alternance de radical.

Les noms massifs tels que *o* 'eau', *sabl* 'sable', *mobilje* 'mobilier', etc. ainsi que les abstraits comme *liberte* 'liberté', *egzaktityd* 'exactitude', etc. sont censés échapper au principe arithmétique puisque leur référent ne consiste pas en un ensemble dénombrable d'exemplaires. Alors que le pluralisation d'un nom comptable ne fait le plus souvent que multiplier les individus référés, il est bien connu que la même opération appliquée à un massif entraîne des altérations sémantiques diverses :  $le=z=o$  'les eaux' dénote ou bien plusieurs sortes d'eaux (du robinet, ferrugineuse, etc.), ou bien une quantité remarquable dudit liquide :  $le=z=o d\grave{a}=Versaj$  'les eaux de Versailles', ou bien le placenta ( $perdr\grave{a} le=z=o$  'perdre les eaux'), et peut-être d'autres choses encore...  $Le=liberte$  'les libertés' concernent des

domaines particuliers. Dans une certaine mesure, on peut considérer que pluraliser un massif revient à le changer en un nom comptable (cf. Nicolas 2002). Je multiplie les restrictions, car, si être abstrait signifie ne pas dénoter des objets matériels auxquels on peut se cogner, cela ne suffit pas à entraîner la massivité : *mɔrfoloji* ‘morphologie’ est un nom abstrait — à qui a-t-elle jamais causé des bleus ? — mais il en existe plusieurs sortes qu’on peut dénombrer aussi aisément que des choses. Inversement, bien des noms évidemment comptables se laissent utiliser dans une acception massive, avec là aussi des effets de sens plus ou moins notables : cf. *mãʒe dy-fjẽ* ‘manger du chien’ comparé à *mãʒe ẽ=fjẽ* ‘manger un chien’ (cf. Kwon & Zribi-Hertz 2006). Les énoncés génériques visant la totalité de l’espèce naturelle — donc nécessairement un nombre d’individus supérieur à un, l’espèce se laissant elle-même concevoir comme un individu abstrait — constituent un contexte fréquent de massification : cf. *Lə=tiranozɔr vive ãn=Amerik-du-Nɔr* ‘Le tyrannosaure vivait en Amérique du Nord’.<sup>104</sup> Pour importants qu’ils soient, ces phénomènes n’en excèdent pas moins les limites de cet ouvrage comme de mes compétences. Je les laisse donc de côté. Du point de vue morphologique, qui nous importe, les deux paradigmes ci-dessus s’appliquent en tout cas aux massifs comme aux comptables, quel qu’en soit le type. Et l’on remarquera que la « conversion » de comptable à massif n’annule pas l’obligation d’un déterminant : *dy=fjẽ* ‘du chien’ (que l’on mange), *de=la=tomat* ‘de la tomate’ (achetée au marché), etc. sont définis. C’est là, semble-t-il, une spécificité du français, FP comme FO, liée sans doute à la présence nécessaire du déterminant partitif. Remarque qu’on peut sans doute étendre à l’ensemble des langues d’oïl (cf., par exemple, Descusses 1986 : 42-43) et des langues d’oc (cf., par exemple, Salvat 1973 : 21-22), soit à tout le gallo-roman. On comparera l’anglais *eat dog*, le portugais *comer cão*, etc.

Les deux incarnations du français n’en divergent pas moins en cela qu’en FO, dans les constructions particulières qui admettent des noms sans déterminant, ceux-ci ne peuvent être non-spécifiés pour le nombre du fait de la flexion synthétique en <-s>. Ainsi, dans les proverbes tels que « Chien qui aboie ne mord pas », les locutions comme « faire cadeau », les syntagmes du type de « cageot de pommes » et les énumérations telles que « veaux, vaches, cochons, couvées », <chien> et <cadeau> sont singuliers parce qu’en opposition privative avec <chiens> et <cadeaux>, ouvertement pluriels, tandis que <pommes>, <veaux>, <vaches>, etc. sont pluriels parce qu’en opposition avec <pomme>, <veau>, <vache>, etc., singuliers par défaut. L’absence de déterminant n’a pour effet que d’induire une interprétation générique (ni définie ni indéfinie) dans le proverbe, le syntagme et l’énumération, et de faire de « faire cadeau » un verbe complexe, équivalent du dérivé franco-africain *cadeauter*. Le FP connaît les mêmes constructions. Mais là, l’absence de déterminant fait que *fjẽ*, *kado*, *pɔm*, *vo*, *vaf*, etc. ne s’opposent à rien pour le nombre comme pour la définitude. Génériques comme en FO, ils sont donc en outre non-spécifiés quant au nombre, ni singuliers, ni pluriels, à la façon de l’haïtien *kay* mentionné plus haut. On en conclut que le FP, à la différence du FO, accepte, mais dans des contextes restreints, les noms comptables nus limités à leur signification lexicale.

## 7. La flexion adjectivale

Elle concerne deux propriétés : le nombre et le genre grammatical. La flexion du premier est semi-analytique ou supplétive, celle du second supplétive. Le français (toutes variétés) ne connaît pas de flexion comparative synthétique telle qu’elle existe en latin, p.ex. *doctus* ‘savant’, *doctior* ‘plus savant’, mais plutôt une flexion analytique pour autant que *ply=savã*

<sup>104</sup> Le pluriel *Le=tiranozɔr vive...* ‘Les tyrannosaures vivaient’ convient également. L’expression de la généralité étant parasite de celle de la définitude, le risque d’ambiguïté est souvent réel : cf. *Lə=fa mjol* ‘Le chat miaule’ », défini ou générique.

‘plus savant’ se laisse analyser comme une périphrase. Les deux exceptions, *bḥ* ‘bon’ / *mējær* ‘meilleur’ et *movε* ‘mauvais’ / *pir* ‘pire’, manifestent une flexion supplétive, quasi imposée pour le premier (?*ply*=*bḥ* ‘plus bon’), facultative pour le second (*ply*=*movε* ‘plus mauvais’). L’ancien français en comptait davantage, p.ex. *granz* ‘grand’ / *graignor* ‘plus grand’, *petiz* ‘petit’ / *mendre* ‘plus petit’, *beaus* ‘beau’ / *belisor* ‘plus beau’, etc. (Moignet 1988 : 31). Sauf nos deux survivants et, petitement, *greigneur*, ils étaient tous sortis de l’usage dès le moyen français (Gougenheim 1951 : 58).<sup>105</sup>

Le FP connaît deux sortes d’adjectifs : variables et invariables. Les seconds sont les plus nombreux, qui ne varient ni pour le nombre ni pour le genre grammatical : cf.  $\tilde{\epsilon}$ =*fa\_nwar* | *pæle* | *sēpatik* ‘un chat noir | pelé | sympathique’ / *de=fa\_nwar* | *pæle* | *sēpatik* ‘des chats noirs | pelés | sympathiques’, *yn=fat\_nwar* | *pæle* | *sēpatik* ‘une chatte noire | pelée | sympathique’ / *de=fat\_nwar* | *pæle* | *sēpatik* ‘des chattes noires | pelées | sympathiques’. Les premiers sont eux-mêmes de deux sortes : ceux qui ne varient que pour le genre, p.ex.  $\tilde{\epsilon}$ =*fa\_blā* ‘un chat blanc’ / *de=fa\_blā* ‘des chats blancs’, *yn=fat\_blāf* ‘une chatte blanche’ / *de=fat\_blāf* ‘des chattes blanches’ ; et ceux qui ne varient que pour le nombre, et seulement au masculin, p.ex.  $\tilde{\epsilon}$ =*fa\_zenjal* ‘un chat génial’ / *de=fa\_zenjo* ‘des chats géniaux’, *yn=fat\_zenjal* ‘une chatte géniale’ / *de=fat\_zenjal* ‘des chattes géniales’, *se=fat\_sḍ\_zenjal* ‘ces chattes sont géniales’. Aucun adjectif ne se fléchit quant aux deux propriétés, à l’exception des adjectifs du type de *zenjal* dont la forme en /-o/ cumule genre grammatical (masculin) et nombre (pluriel). On dira donc plus exactement qu’aucun adjectif en FP ne se fléchit *distinctement* pour les deux propriétés, différence significative d’avec le FO où <géniales> permet de reconnaître deux exposants graphiques : <e> pour le genre grammatical, <s> pour le nombre.

Les adjectifs invariables constituent un ensemble ouvert et productif, les variables un ensemble fermé, sauf pour ceux du type de *zenjal*.

### 7.1. La flexion de nombre

Elle nous retiendra peu, car les adjectifs ne se séparent pas des noms sur ce point. L’invariabilité représente la propriété par défaut. S’écarte du défaut un ensemble assez fourni et productif d’adjectifs caractérisés par une terminaison *-al* au singulier, qui connaissent une flexion de pluriel supplétive, identique à celle des noms de même forme — *zenjal* ‘génial’ / *zenjo* ‘géniaux’, comme *zurnal* ‘journal’ / *zurno* ‘journaux’ —, qui ne s’applique qu’en accord avec des noms masculins comme le montrent les exemples ci-dessus. Cette restriction est sans objet en FO, puisque les adjectifs en <-al> y forment un féminin en <-al<e>>, qui les soustrait à l’alternance <-al> / <-au<x>>. En FP, les adjectifs en *-al* ne se fléchissent pas quant au genre grammatical, d’où la nécessité de prendre en compte le genre inhérent du nom modifié. Quant à la productivité de la formation, des adjectifs tels que *āvirṇāmātal* ‘environnemental’ / *āvirṇāmāto* ‘environnementaux’, *libidinal* ‘libidinal’ / *libidino* ‘libidinaux’ etc. en témoignent.

La phonologie de la terminaison ne suffit toutefois pas à définir la classe des adjectifs à flexion supplétive *-al/-o* en contexte masculin. Comme avec les noms, il y a des exceptions : *bisefal* ‘bicéphale’, *kanibal* ‘cannibale’, *sal* ‘sale’, *fatal* ‘fatal’, *naval* ‘naval’, etc. sont invariables. On peut rendre compte des trois premiers types d’exception par le fait que la finale n’y a pas valeur dérivationnelle comme dans *zenjal* ‘génial’ (cf. *zeni* ‘génie’) et *āvirṇāmātal* ‘environnemental’ (cf. *āvirṇāmā* ‘environnement’) (cf. Bonami & Boyé 2005). Mais ce critère est fragile : *zeneral* ‘général’ / *zenero* ‘généraux’ se fléchit en nombre, mais la base /zener-/ dont on devrait le dériver n’a pas d’existence effective. Il n’est donc pas différent de *sal* ‘sale’, dérivé de rien, ni de *fatal* ‘fatal’ et *naval* ‘naval’ qui, directement

<sup>105</sup> *Mwēdr* ‘moindre’ n’est plus le comparatif de *pəti* ‘petit’ : cf. ??*mari ε mwēdr k\_an* ‘Marie est moindre qu’Anne’. On ne l’emploie plus guère qu’au sens de « moins important », « moins nombreux ».

empruntés du latin classique *fatalis* et *naualis*, ne permettent pas de reconnaître une base (?/fat-/ , ?/nav-/).

En fait, c'est la productivité qui paraît décisive. Tout nouvel adjectif en *-al*, par exemple le récent *manazerjal* 'managérial', fait un pluriel en *-o*. Cela fait de l'alternance *-al/-o* le cas normal, par défaut, tandis que les adjectifs terminés en /al/ invariables en contexte masculin constituent une liste fermée à mémoriser. D'où, certainement, les fréquentes régularisations où la grammaire scolaire voit des fautes et qui, du reste, vont, semble-t-il, dans les deux sens : de l'invariable exceptionnel vers le défaut du type (*de=zã bano* 'des gens banaux'), du type minoritaire vers le défaut général (*de=problem ãvironãmãtal* 'des problèmes environnementaux').<sup>106</sup> Pour quelques unités, l'hésitation est courante et plus ou moins acceptée : cf. *de=prosesys tonal* ~ *tono* 'des processus tonals ~ tonaux'. Les dictionnaires s'efforcent au possible de particulariser l'emploi de l'une ou l'autre forme : les paletots de Rimbaud seraient probablement devenus idéaux ; des compagnons de voyage peuvent être idéals. Je ne suis pas sûr qu'il n'y ait pas là quelque artifice. (Le domaine de l'exception est plus limité en FO, là encore parce que les adjectifs du type de <bicéphale> ne s'y terminent pas en <al>. Parce qu'aucune dérivation ne s'y laisse reconnaître et aussi sans doute sous l'influence de l'orthographe, ils ne sont jamais pluralisés en *-o* : cf. *de=festẽ \*kanibo* 'des festins \*cannibaux', ou alors par humour potache, dont il ne convient pas de mésestimer l'importance, pour ce qu'il révèle du sentiment linguistique.)

## 7.2. La flexion de genre grammatical

Comme le montre l'exemple de *blã / blãf* et abstraction faite des adjectifs en *-al* variables, les adjectifs qui marquent le genre grammatical le font par un procédé où l'on peut voir un supplétisme partiel analogue à celui qui unit *bœf* à *bø*. En fait, il convient de distinguer deux groupes : (a) les paires telles que *bo* 'beau' / *bœl* 'belle', *fu* 'fou' / *føl* 'folle', *liberatør* 'libérateur' / *liberatrīs* 'libératrice', *nœf* 'neuf' / *nœv* 'neuve', *nuvo* 'nouveau' / *nuvel* 'nouvelle', *pasif* 'passif' / *pasiv* 'passive', *rijær* 'rieur' / *rijøz* 'rieuse', *sek* 'sec' / *sef* 'sèche', *vjø* 'vieux' / *vjej* 'vieille', etc. ; (b) les paires telles que *blã / blãf*, *grã* 'grand' / *grãd* 'grande', *pæti* 'petit' / *pætīt* 'petite', etc. Les membres des deux groupes ont pour point commun de se présenter sous deux formes. Ils se distinguent par leur façon de différer. Dans le groupe (b), les deux formes peuvent être dites respectivement courte et longue, la longue, affectée à l'expression du féminin, comptant une consonne finale — parfois deux : cf. *syspẽ* 'suspect' / *syspekt* 'suspecte' — qui fait défaut à la courte (cf. Blanche-Benveniste & Chervel 1969/1978 ; Bonami & Boyé 2005 ; Bonami, Boyé & Tseng 2005). Les formes appariées du groupe (a) ou bien ne diffèrent pas en longueur mais par la nature de leur segment ou syllabe finale (*liberatør / liberatrīs*, *pasif / pasiv*, *rijær / rijøz*, *sek / sef*), ou bien diffèrent en longueur, mais aussi par la qualité de la voyelle finale ou unique du masculin (*bo / bœl*, *fu / føl*, *nuvo / nuvel*, *vjø / vjej*). C'est pour eux que je parlerai d'adjectifs supplétifs. Les membres du groupe (b), je les appellerai simplement biformes. La consonne additionnelle de la forme longue est dite « instable » ou « latente », étiquettes que leur soubassement théorique rend également critiquables. C'est la forme féminine des adjectifs des groupes (a) et (b) qui sert le plus souvent de base de dérivation : *føl.i* 'folie', *pasiv.ite* 'passivité', *sef.war* 'séchoir', *vjej.es* 'vieillesse'. Pour la langue contemporaine *bo.te* 'beauté' et *nuvo.te* 'nouveauté' peuvent être vus comme des exceptions — mais cf. *bœl.atr* 'bellâtre', *nuvel.mã* 'nouvellement'.

S'agissant de l'association à l'expression du genre féminin de la forme longue et/ou à finale modifiée, j'ai exposé au chapitre 3 les raisons qui me font rejeter l'hypothèse sauveuse d'orthographe qui fait de la forme longue la forme « sous-jacente » d'où la courte est dérivée. Forme courte et forme longue existent en parallèle, la première ne résulte pas du

<sup>106</sup> *Bano* 'banaux' est réputé « correct » au sens médiéval de « collectif » (« fours banaux »). Mais qui le sait encore ?

racourcissement de la seconde — ni, à la rigueur, la seconde de l’allongement de la première. Dans ces conditions, la notion de consonne « latente » n’est guère adéquate, car elle implique de fait une relation de ce type, comme le fait aussi, me semble-t-il, quoique moins nécessairement, l’analyse de Bonami, Boyé & Tseng (2005), selon laquelle la représentation phonologique des adjectifs biformes (et aussi des noms comme *fa* ‘chat’ / *fat* ‘chatte’) se compose de deux listes de segments : le corps et l’appendice. La grammaire spécifie les contraintes selon lesquelles l’appendice se réalise ou non. Les adjectifs uniformes sont ceux dont l’appendice est vide. Mais, puisqu’il s’agit de flexion, cette représentation phonologique n’a de réalité qu’au niveau des mots-formes spécifiés pour le trait de genre, tandis que les lexèmes BLANC, CHAT, etc., tout virtuels comme on sait, ne le sont pas. Autrement dit, BLANC diffère de NOIR du fait d’être réalisé par deux mots-formes, masculin et féminin, au lieu de l’être par un seul, épïcène ; et CHAT diffère de GIRAFE du fait d’être exprimé par deux mots-formes, tandis que GIRAFE l’est par un seul, grammaticalement féminin mais référant indifféremment aux deux sexes.

Pourquoi, s’il en est ainsi, ne pas opter plus simplement pour une corrélation directe entre les différences phonologique et sémantique, entre /blā/ et masculin, /blāf/ et féminin, comme le font sûrement les jeunes francophones qui n’ont pour se guider que ce qu’elles et ils entendent et donc guère de moyens de restituer un appendice tantôt audible, tantôt non, et de forme toujours variable ? C’est du reste l’analyse à laquelle se rangent Bonami & Boyé (2005), qui considèrent /blā/ et /blāf/ comme deux « thèmes » qui, en tant que tels, n’expriment aucun genre grammatical en FP contemporain, mais se trouvent y être associés aux deux genres que connaît la langue.<sup>107</sup>

Le coût de cette renonciation à des formes complexes est minime, vu que, de toute façon, les paires masculin-féminin telles que *blā* / *blāf*, *grā* / *grād*, etc. devront être mémorisées une à une, solution réputée peu « élégante » en regard de la règle générale que permet l’hypothèse de la troncation. Mais celle-ci a sa propre inélégance, on l’a vu, à savoir de supposer une entité, le fameux « *e* muet », sans véritable nécessité.<sup>108</sup> On aurait tort, en outre, de sous-estimer les capacités de mémorisation des jeunes humains. D’autant que ces paires constituent un ensemble fermé, car les formations productives donnent ou bien des adjectifs invariables (p.ex. *-abl*, *-ist*), ou bien des paires où la différence est dans une certaine mesure prévisible comme avec *liberataer* / *liberatrix*, *rijær* / *rijæz*, *pasif* / *pasiv*, *ʒwisif* ‘jouissif’ / *ʒwisiv* ‘jouissive’, etc. (voir plus loin).

Pour les adjectifs antéposables, la forme affectée au féminin — et non la forme « féminine » — apparaît en outre devant un nom masculin à initiale vocalique :  $\tilde{\epsilon}=bel\_om$  ‘un bel homme’ ( $yn=bel\_fam$  ‘une belle femme’),  $\tilde{\epsilon}=nuvel\_epizod$  ‘un nouvel épisode’ ( $yn=nuvel\_seri$  ‘une nouvelle série’),  $\tilde{\epsilon}=vjej\_ami$  ‘un vieil ami’ ( $yn=vjej\_ami$  ‘une vieille amie’),  $\tilde{\epsilon}=pätit\_ami$  ‘un petit ami’ ( $yn=pätit\_ami$  ‘une petite amie’).<sup>109</sup> Cela ne concerne pas les adjectifs dont la consonne finale est fixe, quoique éventuellement mutable, comme *sek* / *sef* ‘sec / sèche’ :  $\tilde{\epsilon}=tr\epsilon=sek\_akøj$  ‘un très sec accueil’, non  $*\tilde{\epsilon}=tr\epsilon=sef\_akøj$  ‘un très sèche accueil’ — mais l’antéposition de ces adjectifs est rarement heureuse.<sup>110</sup> Quant au dernier

<sup>107</sup> Pour les auteurs, cette vacuité sémantique inhérente fait des deux thèmes des morphèmes (cf. chapitre 5).

<sup>108</sup> L’important n’est en effet pas tant que le *e* muet n’ait pas d’existence matérielle, mais plutôt que son existence virtuelle n’est pas nécessaire pour rendre compte du phénomène. Certes, l’analyse par alternance de formes est moins familière, elle éloigne le FP des autres langues romanes. Une recherche objective ne saurait toutefois s’arrêter à cet argument.

<sup>109</sup> Il s’agit là, cela va de soi, d’une analyse synchronique. En ancien français, le masculin *novel* [no’veʎ] représentait le cas régime singulier ou le cas sujet pluriel. Devant le *-s* du cas sujet singulier ou du cas régime pluriel, le [ʎ] se vélarisait, d’où *noveaus* [no’vʝaws] ~ [no’vʝos]. Le féminin était *novelle(s)* [no’veʎə(s)] (cf. Raynaud de Lage : 1967 : 43).

<sup>110</sup> Pour une raison qui m’échappe, mais où le poids phonétique doit entrer pour une part, l’antéposition de *sek* sonne mieux quand l’adjectif est amplifié par un quantifieur, qui améliore aussi  $\tilde{\epsilon}=tr\epsilon=sek\_ete$  ‘un très sec été’ comparé au quasi inacceptable  $?\tilde{\epsilon}=sek\_ete$  ‘un sec été’.

exemple, qui met en jeu un adjectif du groupe (b), on parle d'ordinaire de liaison, terme qu'on évite avec le groupe (a). On voit pourtant qu'il s'agit du même phénomène dans la langue contemporaine par suite de la chute des schwas finaux qui laisse dans tous les cas une consonne finale au contact d'une voyelle initiale. Avec les adjectifs biformes terminés en *-ø* comme *kyrjø* 'curieux', on pense aussi avoir affaire à une liaison sur la forme masculine dans /ẽkyrjøzãzẽ/ 'un curieux engin'. Mais c'est sans doute qu'on part de l'orthographe, alors qu'on peut très bien voir là aussi le féminin *kyrjöz* 'curieuse', si bien que  $\tilde{\epsilon}=\text{kyrjöz\_ãzẽ}$  serait la bonne transcription.

L'explication qu'on donne traditionnellement de ce phénomène est qu'il s'agit d'éviter l'hiatus que produiraient /ẽboəm/ et /ẽpøtiami/. Outre les arguments synchroniques qui l'infirmement, qui seront développés plus loin, elle ne tient pas historiquement s'agissant du groupe (a) : Gougenheim (1951 : 53) montre qu'au seizième siècle la répartition des deux formes était souvent aléatoire et ne tenait pas compte de l'initiale du mot suivant (*un beau exemple, un bel spectacle, un vieux homme*) — tandis que la consonne des adjectifs du groupe (b), muette devant consonne, était prononcée devant voyelle, mais aussi devant une pause, où la question de l'hiatus ne se pose pas : p.ex. *Il est petit* /ilɛpøtit/. (Il s'ensuit qu'en moyen français, la consonne peut effectivement être dite « latente ». C'est la disparition de la condition « pause » qui a mené à la situation actuelle.) Le wallon présente grosso modo les mêmes règles que le moyen français (cf. Morin 2005).

Il n'est pas exact, d'autre part, que la sélection de la forme affectée au féminin soit systématique devant voyelle initiale d'un nom masculin en FP actuel. Avec le groupe (a), on observe une variation selon que la suite adjectif-nom est plus ou moins attendue (lexicalisée) :  $\tilde{\epsilon}=\text{vjø}(z)\_ami$  est sans doute plus rare que  $\tilde{\epsilon}=\text{vjej\_ami}$  'un vieil ami' (qui n'est pas forcément âgé), quoiqu'il me paraisse exagéré de l'exclure ; en revanche, (*Nu=søm\_mõte dãz=*) $\tilde{\epsilon}=\text{vjø}(z)\_avjõ$  '(Nous sommes montés dans) un vieux avion', préférablement sans liaison (voir plus loin), paraît assez banal, tout autant que  $\tilde{\epsilon}=\text{nuvo\_epizød dã=la=seri}$  'un nouveau épisode de la série'. La coordination d'adjectifs antéposés suscite des choses comme  $\tilde{\epsilon}=\text{vjej\_ε\_farmã}(t)\_ami$  'un vieil et charmant ami' où la forme pré-vocalique ne précède pas immédiatement le nom qui la requiert et n'est pas due non plus à la conjonction de coordination  $\epsilon$  'et' : cf.  $\tilde{\epsilon}n=\text{ami\_vjø\_ε\_farmã}$  'un ami vieux et charmant' plutôt que ?  $\tilde{\epsilon}n=\text{ami\_vjej\_ε\_farmã}$  ? 'un ami vieil et charmant' (cf. Bonami & Boyé 2013). La structure syntaxique est donc au moins aussi importante que le contexte phonologique, sans oublier le registre, car l'exemple ci-dessus est manifestement affecté ou ludique. Quant au groupe (b), qui tolère plus ou moins  $\tilde{\epsilon}=\text{pøti\_ami}$ , j'y reviens plus loin.

A l'exception des adjectifs en *-if* /*-iv* et *-ær* / *-øz*, où l'alternance est générale, les paires du groupe (a) doivent sans doute être mémorisées comme signalé plus haut, car seul le recours à la diachronie permet de justifier la différence formelle, à savoir qu'il y eut une étape dans l'évolution du français où le lexème FOU avait pour masculin (cas régime singulier ou cas sujet pluriel) /fɔl/ et pour féminin /fɔlə/ ; et une autre où le féminin de *søk* avait pour structure morphologique /søk-ə/, le /ə/ final issu d'un /a/ latin (cf. latin tardif *\*siccu* / *\*sicca*) devant lequel /k/ s'est régulièrement palatalisé (cf. *cheval* /ʃəval/ < *\*ka'βalu*/), processus tôt éteint en ancien français (dès avant les premières attestations) et, par là, peu représenté dans le cas qui nous occupe, en fait limité à cet item, puisque les 128 autres adjectifs à /k/ final prononcé (*c* orthographique) sont tous invariables — cf. *fik* 'chic', *tyrk* 'turc, turque', etc. — tandis que le <c> final de *blã* 'blanc' et *frã* 'franc' n'est jamais réalisé tel quel, à savoir /k/ (voir plus loin).<sup>111</sup>

Les choses sont plus complexes pour les adjectifs comme *liberatær* / *liberatrīs* — cf. *yn=møzr\\_liberatrīs* 'une mesure libératrice'. On a vu au chapitre précédent, à propos des noms de même forme, qu'il y a bien là un fait de flexion. L'alternance serait prévisible n'était

<sup>111</sup> Je remercie Olivier Bonami de m'avoir communiqué le corpus des 131 adjectifs à *c* final.

l'existence d'items terminés en *-tær* au masculin, mais dont le *t* appartient non au suffixe, mais à la base du verbe associé, et dont le féminin est en *-øz* : p.ex. *fāt-e* 'chanter' / *fāt.ær* 'chanteur' / *fāt.øz* 'chanteuse',<sup>112</sup> *flat-e* 'flatter' / *flat.ær* 'flatteur' / *flat.øz* 'flatteuse', *sot-e* 'sauter' / *sot.ær* 'sauteur' / *sot.øz* 'sauteuse' (cf. *yn=pys\_sot.øz* 'une puce sauteuse'), etc. (On ne citera que pour mémoire *āfāt.ær* 'enchanteur' dont le féminin *āfāt.æres* 'enchanteresse' se voit souvent remplacé par *āfāt.øz* 'enchanteuse'.) Aller chercher le verbe base pour prédire la forme de l'adjectif, voilà une opération qui ne rebute pas les linguistes, mais à quoi ne se livrent sûrement pas les enfants acquérant la langue. D'autant que la recherche de la base ne va pas de soi lorsque celle-ci n'est mobilisée qu'à fin de dérivation : cf. *protek.tær* 'protecteur' / *protek.tris* 'protectrice' vs. *protez-e* 'protéger', *destryk.tær* 'destructeur' / *destryk.tris* 'destructrice' vs. *detruji-r* 'détruire', etc. Pourquoi pas *\*protezær* / *\*protezøz* et *\*detruizær* / *\*detruizøz* ? Ces mots sont-ils vraiment inattestés ? Quoi qu'il en soit, on conviendra que ces paires doivent elles aussi être mémorisées au cours de l'acquisition. Ce qui, de même que pour les noms, ne s'oppose pas à ce que leur nombre et leur relative régularité permettent de nouvelles formations par analogie : cf. *avjatær* 'aviateur' (1863 selon le TLFi) dont la féminisation *avjatrīs* 'aviatrice' — sûrement plus tardive — est due à un *t* dont on ne peut décider s'il appartient au suffixe ou à la base, les bases /avja-/ et /avjat-/ étant aussi virtuelles l'une que l'autre.<sup>113</sup>

Les membres du groupe (b) ont aussi pour caractéristique que la consonne finale de la forme longue ne se laisse pas prédire à partir de la forme courte : cf. *blā* 'blanc' / *blāf* 'blanche', *fo* 'faux' / *fos* 'fausse', *frā* 'franc' / *frāf* 'franche', *kokε* 'coquet' / *koket* 'coquette', *leze* 'léger' / *lezer* 'légère', *lō* 'long' / *lōg* 'longue', *prō* 'prompt' / *prōpt* 'prompte', *salō* 'salop, salaud' / *salop* 'salope',<sup>114</sup> *sur* 'sourd' / *surd* 'sourde', *ver* 'vert' / *vert* 'verte', *zalu* 'jaloux' / *zaluz* 'jalouse', etc. Tout aussi imprévisible est l'existence même de deux formes : cf. *zāti* 'gentil' / *zātij* 'gentille', mais *zoli* 'jolie' est invariable.

Quelques généralisations sont néanmoins possibles. Si la forme courte se termine par une voyelle nasale, la forme longue se termine par la voyelle orale correspondante suivie de /n/ : *amerikē* 'américain' / *ameriken* 'américaine', *bō* 'bon' / *bōn* 'bonne', *persā* 'persan' / *persan* 'persane'. Mais cette « règle » souffre d'exceptions, p.ex. *grā* / *grād*, *lō* / *lōg*, *persā* 'perçant' / *persāt* 'perçante' — de même tous les adjectifs en /ā/ associables à un participe présent ou gérondif (cf. chapitre 9) — *sē* 'saint' / *sēt* 'sainte' (mais *sē* 'sain' / *sēn* 'saine'). Le timbre de la voyelle orale ne correspond pas toujours à celui de la nasale : cf., *brē* 'brun' / *bryn* 'brune' (dans ma prononciation, apparemment majoritaire), *divē* 'divin' / *divin* 'divine', *malē* 'malin' / *malij* 'maligne' (où la consonne nasale est imprévisiblement palatale).

Sans exception, en revanche, est la correspondance entre forme courte en /jē/ et forme longue en /jen/ : *āsje* 'ancien' / *āsjen* 'ancienne', *parizjē* 'parisien' / *parizjen* 'parisienne', etc. ; ou entre forme courte en /ø/ et forme longue en /øz/ : *krø* 'creux' / *krøz* 'creuse', *lyminø* 'lumineux' / *lyminøz* 'lumineuse', etc. Il est remarquable que *blø* 'bleu.e', bien que faisant exception car épïcène, possède, sûrement par analogie, une forme longue /bløz/ qui sert de base de dérivation dans *bløzaj* 'bleusaille'.

Comme on l'a vu, deux contextes outre le féminin suscitent la consonne « latente » de la forme longue : la dérivation d'un nom, d'un verbe ou d'un adverbe — cf. *a.surd.ir* 'assourdir', *blāf.ær* 'blancheur', *fos.te* 'fausseté', *fos.mā* 'faussement', *frāf.iz* 'franchise', *krøz.e* 'creuser', *lōg.ær* 'longueur', *zātij.es* 'gentillesse', etc. —<sup>115</sup> et l'antéposition à un nom masculin à initiale vocalique comme dans *ē=petit\_ami pāti* / *ēpətitamipəti* / 'un petit ami petit'.

<sup>112</sup> Cf. *ēn=wazo\_fātær* 'un oiseau chanteur', *de=ras...\_fātøz* 'des races... chanteuses' (TLFi).

<sup>113</sup> *Avjatær* a aussi un emploi adjectival (cf. TLFi). D'une façon générale, faut-il le préciser ? les adjectifs en *-tær* / *-tris* sont aussi des noms. Mais l'inverse n'est pas vrai : bien des noms en *-tær* / *-tris* ne s'emploient pas comme adjectifs. Quant à l'ancienneté d'*avjatrīs*, mon seul indice à ce jour est que la première femme à obtenir un brevet de pilote fut Raymonde de Laroche en 1910.

<sup>114</sup> Le féminin *salod* 'salaude' est obsolète (cf. TLFi).

Le SN ainsi formé doit être singulier, sinon la consonne n'apparaît pas : cf. *de=pəti\_z=ami* /depətizami/ 'des petits amis' — comme on pouvait s'y attendre, l'interposition de *z* faisant que l'adjectif n'est pas alors en contexte prévocalique. (Elle est susceptible d'apparaître, en revanche, si *z* ne s'insère pas — voir plus loin.)

La consonne manifestée est souvent la même dans les trois contextes (cf. *pətit\_kopin* 'petite copine', *pətit.es* 'petitesse'). Mais il y a des exceptions : la consonne « latente » de *grā* 'grand' est /d/ au féminin et en contexte de dérivation (*grād*, *grād.œr*), mais /t/ en contexte d'antéposition à un masculin (*ẽ=grāt\_ami* 'un grand ami') ; celle de *gro* est /s/ dans *gros* 'grosse' et *grosist* 'grossiste', mais /z/ dans *groz\_idjo* 'gros idiot'. En l'absence de déterminant, l'opposition de voisement peut constituer le seul indice audible de la différence de genre : cf. *grāt\_ẽbesil* ! 'Grand imbécile !' vs. *grād\_ẽbesil* ! 'Grande imbécile !'. La phonologie, à savoir la position intervocalique de la consonne, n'explique pas ces différences. Qu'une occlusive se dévoise entre voyelles (/d/ → /t/) n'est pas ordinaire, on attend plutôt le contraire comme dans l'espagnol *cada* 'chaque' du grec *kata* ; dans *grosist*, le /s/ est intervocalique, mais ne se voise pas. Il est des cas indécidables. Ainsi, la consonne « latente » de *ver* 'vert' est /t/ au féminin (*vert* 'verte'), mais /d/ en dérivation (*verdir* 'verdier', *verdyr* 'verdure').<sup>116</sup> Mais *ver*, comme les autres adjectifs de couleur en discours ordinaire, ne précède pour ainsi dire jamais le nom qu'il qualifie, si bien qu'on manque d'exemples attestés pour être assuré de la consonne, /t/ ou /d/, qui s'entendrait dans, par exemple (pardon Baudelaire), *lə=ver(?)\_eden de=z=amur\_(z=)ãfãtin* 'le vert éden des amours enfantines'. Probablement /t/ par analogie. Mais le plus probable est qu'on éviterait toute liaison et s'en tiendrait à l'enchaînement avec la consonne fixe : /vɛkɛdɛn/. Cette incertitude peut même conduire à empêcher devant voyelle une antéposition par ailleurs possible : *ẽ=blō\_nɔrdik* 'un blond Nordique' passe encore, mais on recule devant *?ẽ=blō( )\_ãfã* 'un blond enfant', indécis qu'on est quant à la consonne qui devrait remplir la parenthèse vide : *blōd\_ãfã* est décidément impossible, *blōt\_ãfã* douteux, *blō\_ãfã* gauche ; on dira donc *ẽn=ãfã blō* 'un enfant blond' (cf. Morin 2005). En revanche, *yn=blōd\_ãfã* 'une blonde enfant' est sans problème.

Pourquoi ces différences ? Les études sur la question, en particulier Bonami & Boyé (2005) et Morin (2005) n'en proposent pas vraiment d'explication, et l'on peut douter qu'il y en ait une, cette complexité résultant d'une suite de processus historiques irrécupérables. Mais elles tâchent d'en élaborer des descriptions explicites basées sur des concepts clairement définis, sans doute le mieux qu'on puisse faire.

Se fondant sur Bonami & Boyé (2005), on peut considérer que le paradigme des adjectifs genrés comporte cinq formes : (a) une forme affectée au masculin (FM) ; (b) une forme affectée au féminin (FF) ; (c) pour les adjectifs en *-al* variable seulement, une forme affectée au masculin pluriel (FMP) ; (d) une forme servant de base dérivationnelle (FBD) ; (e) une forme réservée à la liaison avec un nom masculin singulier (FLMS). Pour autant que le /z/ qui s'entend dans *de=blō\_z=ãfã* 'des blonds enfants' et *de=blōd\_z=ãfã* 'des blondes enfants' n'appartient pas à l'adjectif (cf. chapitre 6), il n'est pas nécessaire de prévoir une forme de pluriel hors du cas (c), dont on fera abstraction. Les rapports entre les quatre bases restantes dépendent de la FM, selon que celle-ci se termine par une voyelle ou une consonne. Dans le premier cas, la FM diffère des trois autres, qui sont identiques entre elles : cf. *vjə\_frer* 'vieux frère' en regard de *vjej\_tāt* 'vieille tante', *vjej\_ðkl* 'vieil oncle', *vjejes* 'vieillesse'. Mais si la FM se termine par une consonne mutable, la FLMS lui est identique : cf. *viv\_emosjð* 'vive émotion', *vivasite* 'vivacité', mais *vif\_emwa* 'vif émoi' comme *vif\_sātīmã* 'vif sentiment' ; *sef\_reparti* 'sèche répartie', *sefərəs* 'sècheresse', mais *trɛ=sek\_akæj* 'très sec accueil' comme *trɛ=sek\_rəfy* 'très sec refus'. (Les adjectifs terminés par une consonne non mutable, par

<sup>115</sup> En français méridional, où l'existence du schwa féminin paraît probable, celui-ci est régulièrement prononcé dans les adverbes en /mã/ : /fosəmã/, *aktɥeləmã* /aktɥeləmã/ 'actuellement', etc.

<sup>116</sup> Le féminin *verde* se rencontre en moyen français (Gougenheim 1951 : 49).



exemple *fik* ‘chic’, *synit* ‘sunnite’, *rapid* ‘rapide’, etc., sont invariables et leur paradigme ne comporte qu’une seule forme.)

Les exceptions à cette généralité sont peu nombreuses. J’ai mentionné plus haut *grā* et *gro*, dont les FLMS *grāt* et *groz* diffèrent de toutes les autres formes. On ajoutera, semblables à *grā*, *profō* ‘profond’ (*profōd\_emosjō* ‘profonde émotion’, *profōdær* ‘profondeur’, mais *profōt\_emwa* ‘profond émoi’) et *sægō* ‘second’ (*sægōd\_vy* ‘seconde vue’, *sægōder* ‘secondaire’, mais *sægōt\_etaz* ‘second étage’) ; semblables à *gro*, *bā* ‘bas’ (*bas\_kalite* ‘basse qualité’, *bases* ‘bassesse’, mais *baz\_etaz* ‘bas étage’), *du* ‘doux’ (*dus\_ami* ‘douce amie’, *dusær* ‘douceur’, mais *duz\_ami* ‘doux ami’) et *fo* ‘faux’ (*fos\_ide* ‘fausse idée’, *fosæte* ‘fausseté’, mais *foz\_ami* ‘faux ami). En tout, sept unités. Les autres candidats possibles parmi les adjectifs éventuellement prénominaux, tels *blō* ‘blond’ et *fo* ‘chaud’, sont douteux, car on évite le plus souvent leurs FLMS théoriques, comme on l’a vu pour *blō*. Et qui oserait  $\tilde{\varepsilon}$ =*fo\_t\_ete* ‘un chaud été’ pour pallier l’impossibilité de  $\tilde{\varepsilon}$ =*fod\_ete* ‘un chaud été’, pourtant conforme à la généralisation ?

On en conclura deux choses. Premièrement, il semble vrai que, si la FF d’un adjectif fléchi pour le genre grammatical se termine par /d/ ou /s/, la FLMS se termine par /t/ ou /z/. Mais, c’est le second point, il semble également que cette alternance soit sentie comme aberrante, si bien qu’elle n’est pratiquée que pour un très petit nombre d’adjectifs dont la position prénominale est naturelle, voire parfois obligatoire, comme dans *grāt\_om* ‘grand homme’ ou *foz\_ami* ‘faux ami’. Partout ailleurs, on fait en sorte d’éviter d’avoir à employer la FLMS théorique, pour autant qu’elle existe, ou bien en ne liant pas sans souci de hiatus ( $\tilde{\varepsilon}$ =*fo\_ete* ‘un chaud été’), ou bien en évitant l’antéposition ( $\tilde{\varepsilon}$ *n=ete\_fo* ‘un été chaud’).

Le modèle de Bonami & Boyé n’exclut pas la notion de consonne « latente », que, du reste, les auteurs adoptent. Mais il est également compatible avec une approche intégralement « surfaciste », *wysiwyg* /ˈwɪzɪwɪɡ/ ‘What you see is what you get’, que je traduirais volontiers à la Queneau — peut-être en en modifiant quelque peu le sens — par *squonvouacesquia* /skōwasɛskja/ ‘Ce qu’on voit, c’est ce qu’il y a’, fondée sur les relations entre des formes indécomposables.

La description de Morin (2005) s’écarte plus radicalement des traitements traditionnels. Constatant que « l’évolution générale du français révèle une absence totale d’effets dus à la contrainte anti-hiatus ... sur la liaison » (p. 17), il avance que le /t/ intermédiaire qu’on entend dans /pətɪtami/ ‘petit ami’ est à rattacher au nom plutôt qu’à l’adjectif. Son hypothèse est que les noms masculins à initiale vocalique sont intériorisés sous plusieurs formes dans le lexique mental des francophones, à savoir une forme d’*état libre* (*ami*) et plusieurs formes d’*état construit* suscitées par certaines séquences : par exemple, *t-ami* après *pəti*, *grā*, etc., *n-ami* après  $\tilde{\varepsilon}$  ‘un’, *āsjē* ‘ancien’, etc., /t/ et /n/ des préfixes.

Etat libre et état construit (aussi dit d’annexion) sont des notions héritées de la grammaire des langues chamito-sémitiques (arabe, berbère, hébreu, etc.). En berbère kabyle, par exemple, le nom voulant dire ‘homme’ est *argaz* à l’état libre, *wergaz* à l’état construit simple lorsqu’il suit le verbe dont il est le sujet ou qu’il est précédé d’une préposition — *argaz iruh* ‘l’homme est parti’ vs. *iruh wergaz* ‘est parti l’homme’, *d wergaz* ‘avec l’homme’ — et à l’état construit renforcé lorsqu’il complète un autre nom — *awal wwegaz* ~ *bbwegaz* ‘la parole d’un homme’ (Mammeri 1986 : 29-31).<sup>117</sup> En berbère, la causalité est syntaxique. En FP, elle serait plutôt lexicale, seules certaines collocations amenant l’état construit : cf. *pəti\_t-ami* vs. *zoli\_ami* ‘joli ami’.

Morin appuie son hypothèse sur les travaux qui montrent que les enfants acquérant le français passent par une étape où elles et ils oscillent entre plusieurs réalisations du même lexème, par exemple /uʁs/, /nuʁs/, /zuʁs/, /tuʁs/ pour OURS (cf. chapitre 6 ; cf. Côté 2005 ;

<sup>117</sup> La plupart des noms masculins commencent par une voyelle /a/ ou /i/. Les noms masculins, plus rares, à initiale consonantique ne connaissent pas ces alternances d’état.

voir aussi Chevrot et al. 2005). Il cite aussi le témoignage d'un Louisianais francophone, mais illettré en français, qui ne voit aucun inconvénient à ce que plusieurs mots, /aʁbʁ/, /laʁbʁ/, /naʁbʁ/, /zaʁbʁ/, correspondent aux deux mots anglais *tree* et *trees*. Voilà qui donne une idée de ce que peut être la compétence d'un francophone adulte exempt de toute influence orthographique. Mais est-ce pertinent pour celle de l'immense majorité des francophones québécois et européens (y compris les DOM-TOM), qui subissent tôt cette influence ?<sup>118</sup> On peut en douter. Quant aux exemples impliquant /z/ — « B : Plusieurs z-années, plutôt. / A : Z-années !?, je ne crois pas. » (Morin 2005 : 11) — ils ne peuvent être pertinents si z est bien un exposant proclitique de la pluralité, sans rapport direct avec la question des collocations adjectif-nom.

En fait, l'hypothèse de Morin ne s'applique raisonnablement qu'aux collocations où l'adjectif prénominal est de ceux auxquels la tradition attribue une consonne « instable ». Peut-être OURS présente-t-il en effet trois états construits : *nurs* dans  $\tilde{\epsilon}$ =*nurs* 'un ours', *turs* dans *pəti\_turs* 'petit ours', *zurs* dans *gro\_zurs* 'gros ours'. Mais va-t-on supposer *kurs* à cause de *fik\_urs* 'chic ours' ? Non, sans doute, et Morin ne le fait pas.

L'indéniable est que les enfants passent par une étape d'incertitude, manifestement due aux difficiles segmentations auxquelles les confrontent les phénomènes de liaison du FP. Mais ils ne tardent pas à détecter les régularités et à ne plus produire ni \* $\tilde{\epsilon}$ =*turs*, ni \* $\tilde{\epsilon}$ =*gro\_nurs* — encore que *nurs* et son hypocoristique *nunurs* 'nounours', plus ou moins lexicalisés, aient une survie assez longue, parfois égale à celle de la locutrice ou du locuteur. Dès lors qu'il en est ainsi, on ne voit plus très bien ce qui permet de décider en faveur de l'une ou de l'autre analyse, Bonami & Boyé ou Morin. Comment savoir ce que nous, francophones, avons dans la tête : *pətit\_urs* ou *pəti\_turs* ? La question a-t-elle un sens ? Peut-être n'avons-nous que /pətiturs/, dont nos connaissances lexicales et morphosyntaxiques nous informent que ce n'est pas un mot, mais un syntagme. S'il me faut vraiment choisir — et je suppose qu'il le faut — j'opterai pour la première, qui présente, me semble-t-il, pour principal avantage que la pluralité des bases réalisant un même lexème est une caractéristique évidente et générale de la flexion du FP (et du FO), alors que l'opposition d'un état libre à un état construit ne concernerait qu'un petit nombre de cas. Le modèle de Bonami & Boyé a donc pour lui sa parcimonie. Attribuant l'alternance des formes à l'adjectif plutôt qu'au nom, il paraît également plus compatible avec l'importante variation qui affecte le phénomène.

La liaison, d'un adjectif biforme antéposé à un nom masculin singulier à initiale vocalique, en principe obligatoire en FO oralisé, est en effet très variable en FP. Ainsi, pour *pəti(t)* 'petit', Mallet (2008 : 324) trouve 75% de liaisons réalisées, et seulement 55,6% pour *grā(d/t)* 'grand'. Divers facteurs expliquent la variation, dont deux principalement. L'un est la nature de la consonne « latente » : /z/ et /t/ sont celles qui se manifestent le plus fréquemment, la première indépendamment de son statut morphologique. On entend  $\tilde{\epsilon}$ =*gro\_āfā*, mais  $\tilde{\epsilon}$ =*groz\_āfā* 'un gros enfant' n'est pas rare — quoique bien sûr moins fréquent que *de=gro\_z=āfā* 'des gros enfants'. En revanche,  $\tilde{\epsilon}$ =*lōg\_iver* s'entend moins que  $\tilde{\epsilon}$ =*lō\_iver* 'un long hiver' — pour ne rien dire de  $\tilde{\epsilon}$ =*lōk\_iver*, aussi archaïque que *sāk\_ēpyr* pour *sā\_ēpyr* 'sang impur'. Ici, le malaise que provoque le dévoisement de /d/ et le voisement de /s/ s'est soldé par un rejet quasi total de ce type de FLMS.

Le deuxième facteur, déjà cité à propos du pluriel et du groupe (a), est le caractère plus ou moins lexicalisé de la suite adjectif-nom. A ce propos, on ne peut que regretter le choix de *pəti(t)* 'petit' comme adjectif type dans beaucoup d'études consacrées à la liaison. Il est vrai que *pətit\_ami* 'petit ami' est la réalisation la plus commune — 75% pour Mallet (2008) — donnant ainsi l'impression que le FP s'aligne sur le FO oralisé. Mais la raison de cette préférence n'est-elle pas plutôt que *pətit\_ami* est un syntagme quasi lexicalisé, une « locution

<sup>118</sup> Je ne me prononce pas à propos des francophones d'Afrique, domaine trop peu exploré qui nous apprendrait beaucoup.

toute faite » comme dit Grammont (1914/1963 : 132), dont le sens n'est clairement pas compositionnel ? De même, plus nettement encore me semble-t-il, *grāt\_om* 'grand homme', rarement, voire jamais prononcé ?/gṛãom/. On est donc justifié à rapprocher ces expressions de locutions indéniablement « toutes faites » et invariablement liées selon Grammont (1914/1963 : 131) telles que *də=baz\_ã=o* 'de bas en haut', *pətit\_a=pəti* 'petit à petit', *pot\_o=fø* 'pot-au-feu', *mot\_a=mo* 'mot à mot', etc. Du reste, depuis 1914, la tendance « pas de liaison » n'a pas non plus épargné cette classe. On dit communément  $\tilde{\epsilon}=pje\_a=ter$  'un pied-à-terre' plutôt que  $\tilde{\epsilon}=pjet\_a=ter$  que sa consonne « latente » réalisée distinguait selon Grammont de l'hiatus de l'expression littérale *mētr pje a=ter* 'mettre pied à terre'. De même, *də=ba\_ã=o* et *də=mjø\_ã=mjø* concurrencent *də=baz\_ã=o* et *də=mjøz\_ã=mjø* 'de mieux en mieux'. *Mot\_a=mo* et *pot\_o=fø* ont résisté, sans doute parce que complètement lexicalisés.

Le problème pour la recherche d'exemples, on l'a vu à propos du pluriel, est que, vu la sémantique associée à l'antéposition des adjectifs (cf. Nølke 1996), il est assez difficile d'en trouver qui n'aient pas peu ou prou le caractère d'une expression lexicalisée. Et *pəti(t)*, parce qu'il tend à fonctionner comme l'équivalent analytique des suffixes diminutifs-hypocoristiques des autres langues romanes — comparez  $\tilde{\epsilon}=pəti\_kafe$  et portugais *um cafézinho* 'un petit café' — est probablement le plus mauvais choix. Le fait que  $\epsilon l=a\ boku\ də=pətit\ ami$  'Elle a beaucoup de petits amis' soit, me semble-t-il, aussi acceptable, voire aussi commun que  $\epsilon l=a\ boku\ də=pəti\_z=ami$  démontre le possible traitement de *petit\_ami* comme un composé plutôt que comme un syntagme. On comparera la quasi impossibilité de \*/dəsavãtãgole/ réalisant l'expression non lexicalisée *de savants Angolais*, seulement prononçable /dəsavããgole/ ou /dəsavãzãgole/. C'est pourquoi j'ai choisi  $\tilde{\epsilon}=l\tilde{z}(g)\_iver$  ci-dessus ;  $\tilde{\epsilon}=gro(z)\_ëk\tilde{z}venjã$  'un gros incon vénient' et  $\tilde{\epsilon}=gro(z)\_wazo$  'un gros oiseau' sont aussi de meilleurs choix.

### 7.3. Pour récapituler

Le cas par défaut des adjectifs en FP est de ne présenter aucune flexion de nombre ni de genre grammatical et donc aucun accord avec les noms qu'ils modifient en tant qu'épithètes ou attributs. C'est le type productif auquel se rattachent les innovations, dérivations nouvelles (*ëförmatic* 'informatique'), emprunts (*kjut* 'cute'), conversions (*il=ε tro* 'Il est trop'). Leur paradigme se réduit à une seule forme — à moins qu'on y ajoute une base de dérivation sur la foi d'alternances telles que *nwarsær* 'noirceur' face à *nwar* 'noir.e(s)', *otomatisite* 'automaticité' face à *otomatic* 'automatique', *sirkylarite* 'circularité' face à *sirkylær* 'circulaire', *kyrjözite* 'curiosité' face à *kyrjöz* 'curieuse', etc. Je ne suis pas convaincu de la nécessité de cette complication. Le cas de *nwarsær* paraît isolé (cf. *ruzær* 'rougeur', *larzær* 'largeur', etc.), et les alternances /k/ → /s/ et /ε/ → /a/, automatiques dans ces formations, peuvent être confiées à la composante dite « morphophonologie » qui assure l'interface entre morphologie et phonologie, autrement dit qui donne une image vocale (pour parler comme Saussure) aux structures morphologiques. (Mais a-t-on même vraiment besoin de ce dispositif ? Après tout, il s'agit en dernière analyse d'une question historique. Par exemple, *kyrjöz(z)* représente l'évolution semi-populaire du latin classique *curiosus*, tandis que *kyrjözite*, plus savant, conserve le timbre de la troisième voyelle de *curiositas*, d'où au final l'alternance /ø/-/o/. Selon le TLFi, les deux mots datent du douzième siècle, l'adjectif de la première moitié, le nom de la seconde. Par analogie, l'alternance s'est étendue aux paires sans ancêtre en latin classique, ainsi *dãzærø* 'dangereux' et *dãzærosite* 'dangerosité'.)

Deux ensembles s'écartent du défaut. D'une part, la majorité des adjectifs de forme *Xal*, invariables quand ils modifient un nom féminin, ont un pluriel *Xo* quand le nom est masculin. Autrement dit, *Xal* est épécène, non spécifié quant au nombre en contexte féminin, singulier en contexte masculin ; *Xo* est masculin pluriel. Un tableau vaut mieux qu'un long discours :

	singulier	pluriel
masculin	Xal	Xo
féminin	Xal	Xal

Les adjectifs du type *Xal/Xo* constituent un ensemble ouvert et productif, si bien que la terminaison s’y laisse identifier comme un formant dérivationnel tel que *Xal* signifie en gros « qui concerne  $\wedge X$  »,  $\wedge X$  un concept exprimé par une base nominale *X* (cf. *āvirɔnəmā* ‘environnement(s)’ ↔ *āvirɔnəmātal* ‘environnemental’). Il existe un sous-ensemble peu fourni et fermé où la terminaison n’est pas ainsi identifiable et dont les membres sont invariables (*fatal* ‘fatal(s)’). Et un sous-sous-ensemble dont les membres, virtuellement décomposables (*ideal* ‘idéal’) ou non (*ostrāl* ‘austral’), acceptent les deux formes.

L’autre ensemble est celui des adjectifs qui se fléchissent pour le genre grammatical, mais pas pour le nombre. C’est un ensemble fermé, mais qui compte beaucoup d’adjectifs basiques, si bien que leur fréquence en discours est grande. Ils sont de deux sortes : supplétifs, dont les formes masculine et féminine sont de longueur égale et peuvent différer en plusieurs points, ainsi *fu* ‘fou’ / *fəl* ‘folle’, *sek* ‘sec’ / *sef* ‘sèche’, *vif* ‘vif’ / *viv* ‘vive’, *vjø* ‘vieux’ / *vjej* ‘vieille’, etc. ; biformes, dont la forme féminine est plus longue d’une consonne que la masculine, ainsi *blā* ‘blanc’ / *blāf* ‘blanche’, *grā* ‘grand’ / *grād* ‘grande’, *gro* ‘gros’ / *gros* ‘grosse’, *lš* ‘long’ / *lšg* ‘longue’, etc. Autant le supplétisme des premiers que la consonne additionnelle des seconds sont imprédictibles. Tous, supplétifs comme biformes, présentent une forme spéciale lorsqu’ils précèdent un nom masculin singulier à initiale vocalique (FLMS). Elle est identique à la forme féminine si la masculine se termine par une voyelle (cf.  $\tilde{\varepsilon}$ =*vjø*\_ *barbš* ‘un vieux barbon’,  $\tilde{\varepsilon}$ =*vjej*\_ *ɔm* ‘un vieil homme’) ; à la masculine si celle-ci se termine par une consonne mutable (cf.  $\tilde{\varepsilon}$ =*vif*\_ *deba* ‘un vif débat’,  $\tilde{\varepsilon}$ =*vif*\_ *emwa* ‘un vif émoi’). Quand la consonne additionnelle d’un adjectif biforme est /d/, /g/ et /s/, l’orthoépique veut que les deux premières se dévoisent (*grāt*\_ *arbr* ‘grand arbre’, *lšk*\_ *iver* ‘long hiver’), tandis que la dernière se voise (*groz*\_ *āfā* ‘gros enfant’). Seuls le dévoisement de /d/ et le voisement de /s/ survivent en FP contemporain.

Divers facteurs contrarient la généralité de cette « règle ». Le premier est sans doute la variation inhérente affectant la liaison, qui ne se fait pratiquement jamais à 100%, si bien que *gro*\_ *āfā* n’a rien d’inouï. S’agissant de l’antéposition des adjectifs, celle-ci n’est vraiment naturelle en FP qu’avec les collocations ayant le caractère ou bien d’un composé (*pətit*\_ *ami* ‘petit ami’), ou bien d’une locution toute faite (*a*=*mš*=*grāt*\_ *etɔnəmā* ‘à mon grand étonnement’). La liaison est alors quasi invariable et sans doute lexicalisée. Dans tout autre cas, on évite en général d’antéposer, en sorte que la FLMS virtuelle n’est presque jamais employée. Il arrive même qu’on ne sache pas trop ce qu’elle pourrait être : dirait-on, selon la règle,  $\tilde{\varepsilon}$ =*blāf*\_ *iver* ? Non, évidemment. Ou bien, flouant la règle, mais se fiant à l’orthographe,  $\tilde{\varepsilon}$ =*blāk*\_ *iver* ? Peu probable. On s’en tient donc à  $\tilde{\varepsilon}$ =*iver*\_ *blā* ‘un hiver blanc’ ou, à la rigueur,  $\tilde{\varepsilon}$ =*blā*\_ *iver*.<sup>119</sup>

Abstraction faite de ces accidents, la flexion « idéale » des adjectifs variables du FP se laisse décrire comme constituant un paradigme à cinq cases, comme le montre le tableau ci-dessous pour quatre unités représentatives :

	FM	FF	FLMS	FBD	FMP
VIF	vif	viv	vif	viv-	
VIEUX	vjø	vjej	vjej	vjej-	

<sup>119</sup> Je choisis à dessein un exemple qui pourrait être une locution à la manière de *blā*\_ *bone* ‘blanc bonnet’. On n’imagine pas d’inverser *urs*\_ *blā* ‘ours blanc’.

GRAND	grã	grãd	grāt	grãd-	
GÉNIA L	zenjal		(zenjal-)	zenjo	

#### 7.4. La question de *kel* ‘quel’

Je réserve pour clôre ce chapitre une place à part à l’adjectif interrogatif-exclamatif *kel* de *kel ε sñn=aʒ ?* ‘Quel est son âge ?’, *kel sñ se=z=ide ?* ‘Quelles sont ses idées’, *kel nə=fy=pa mñ=dezezpwar !* ‘Quel ne fut pas mon désespoir !’. Comme la majorité des adjectifs et comme son alter ego déterminant — cf. *kel=aʒ εl=a ?* ‘Quel âge elle a ?’ — il ne varie ni en genre, ni en nombre. Il présente en outre la propriété de n’apparaître qu’en tête de proposition : cf. *\*sñn=aʒ ε kel?* ‘Son âge est quel?’, en quoi il se distingue du pronom interrogatif *lakel ?* | *lakel ?* | *lekel ?* ‘lequel ? | laquelle ? | lesquels ?’ : cf. *lakel ε sa=mezñ ? ~ sa=mezñ ε lakel ?* ‘Laquelle est sa maison ? ~ ‘Sa maison est laquelle ?’ (voir chapitre 8).

Quoique la question ne concerne pas directement la flexion, on peut néanmoins s’interroger sur cette particularité de *kel?* adjectif. Pourquoi *\*sñn=aʒ ε kel?* ‘Son âge est quel?’ est-elle une phrase agrammaticale, alors que *sñn=aʒ ε(t) avãse* ‘Son âge est avancé’ est parfait, de même du reste que *avãse ε sñn=aʒ* ‘Avancé est son âge’, acceptable en FP quoique un peu littéraire ? La réponse de la GGF est que *kel?*, de par sa nature de mot-*qu*, est affecté de la fonction « extrait », et cela de façon invariable. Autrement dit, *kel?* s’inclut toujours dans la construction interrogative canonique [Q<sub>i</sub>... SX<sub>i</sub>] où il occupe la position initiale distanciée de l’élément qu’il qualifie. Pour les autres adjectifs, non-*qu*, cette fonction est seulement possible, ils peuvent être extraits afin de les mettre en relief (thématisation) comme dans l’exemple ci-dessus. Du fait du caractère obligatoire de sa position initiale, *kel?* ne partage pas cette fonction discursive de thématization : cf. *avãse, sñn=aʒ l=ε* ‘Avancé, son âge l’est’ vs. *\*kel, sñn=aʒ ε•t-il ?* ‘Quel, son âge est-il ?’. Quant à la mobilité de *lakel ?*, elle n’a rien de surprenant : il est bien connu qu’en FP les mots interrogatifs (à l’exception de *kel*) ou bien occupent la position d’extraction liée à un trou syntaxique — [k<sub>i</sub> em=ty e<sub>i</sub>] ‘Qui aimes-tu ?’ — ou bien demeurent dans la position canoniquement associée à leur fonction grammaticale : *ty=em ki ?* ‘Tu aimes qui ?’. (Je reviendrai sur cette question aux chapitre 8 et 14.)

Si le nom que *kel* modifie, dont il est toujours séparé, est pluriel et le successeur vocalique, le proclitique pluriel *z* peut s’insérer : *kel (z=)ete se=z=ide ?* ‘Quelles étaient ses idées ?’ On relèvera néanmoins le parfum littéraire (FO oralisé) de ces expressions, de même que de l’exclamative *kel nə=fy=pa...*, qui ou bien les écarte du FP spontané — on dira plutôt *se=z=ide, s=ete kwa ?* ‘Ses idées, c’était quoi ?’ — ou bien en fait des énoncés ludiques.

### 8. La flexion pronominale

Les pronoms (aussi appelés « proformes ») sont des mots fonctionnels dépourvus de dénotation propre, mais qui tiennent lieu de mots pleins énoncés dans l’environnement discursif proche (antérieur ou à venir), de telle sorte que le locuteur se comporte comme s’il s’estimait en droit de ne pas les répéter *in extenso*, supposant que son auditoire est ou sera capable de les restituer.<sup>120</sup> Le FP connaît trois espèces de pronoms : personnels, démonstratifs (déictiques) et relatifs / interrogatifs / exclamatifs. Il sera surtout question ici de leur forme. On se reportera à Abeillé & Godard (2021, chapitre 9) pour leurs usages.

<sup>120</sup> La formulation est un peu alambiquée, mais comment l’éviter ? Qui sait ce qui se passe dans la tête des locuteurs — y compris la sienne ? De par les suppositions qu’il implique l’emploi des pronoms a aussi une fonction sociale : resserrer les liens du groupe (cf. Goffman 1986).

### 8.1. Les pronoms personnels

Ils se divisent en forts (non clitiques) et faibles (clitiques). Les formes de ces derniers distinguent la personne (1, 2 ou 3), le nombre (singulier ou pluriel), le cas (sujet, objet direct, objet indirect, oblique) et, à la troisième personne, le genre grammatical (masculin, féminin, neutre).<sup>121</sup> Les pronoms personnels forts permettent les mêmes distinctions, sauf le cas. Du point de vue syntaxique, les pronoms personnels faibles expriment les arguments du verbe pivot (tête) de la proposition, tandis que les forts n'entretiennent pas avec lui de relation syntaxique directe. Les pronoms faibles sujet, OD et OI se laissent analyser comme autant de clitiques personnels participant de la flexion verbale. J'y reviens au chapitre 9. Pour les besoins du présent chapitre, je m'en tiendrai à la distinction fort / faible.

Les deux tableaux ci-dessous présentent les paradigmes des pronoms faibles et forts :

	1S G	2S G	3SG	1PL	2PL	3PL
sujet	ʒə	t(y)	i(l) / ε(l) / sə / sa / ʒ̃(n)	ʒ̃(n) ~ nu(z)	vu(z)	i(l) / ε(l)
OD	mə	tə	lə / la / l	nu(z)	vu(z)	le
OI	mə	tə	lɥi	nu(z)	vu(z)	lœr
réfléch i	mə	tə	sə	nu(z)	vu(z)	sə
oblique			i / ā			i/ā

Tableau 8.1 : paradigmes des pronoms personnels faibles

1SG	2S G	3SG	1P L	2P L	3PL	RÉFLECH I
mw a	twa	lɥi / εl	nu	vu	ø / εl	swa

Tableau 8.2 : paradigme des pronoms personnels forts

**8.1.1. Les pronoms personnels faibles** — Leur flexion n'est pas canonique, en ce sens que les contenus de plusieurs cases qui expriment des propriétés distinctes ne diffèrent pas quant à la forme. Aux 1<sup>ère</sup> et 2<sup>e</sup> personnes du singulier, les cases OD, OI et réfléchi sont identiques. Aux 1<sup>ère</sup> et 2<sup>e</sup> personnes du pluriel, abstraction faite de ʒ̃(n), les trois cases sujet, OD et OI ont même contenu.

**8.1.1.1. Les pronoms personnels faibles sujets** — Les pronoms singuliers, 3PL et ʒ̃(n) 'on' pronom 1PL, bien distincts de leurs contreparties fortes (*nu* 'nous' pour ʒ̃(n)), sont clairement des formes clitiques attachées au premier élément du complexe verbal (cf. chapitre 9). Les choses sont moins évidentes avec *nu(z)* 'nous' et *vu(z)* 'vous', qui ne se distinguent de *nu* 'nous' et *vu* 'vous' forts que par le fait de présenter un allomorphe devant voyelle (voir plus loin). Cela me paraît néanmoins suffisant pour les considérer comme des formes clitiques à l'instar des autres pronoms faibles.

La voyelle de la forme 2SG sujet s'élide souvent devant voyelle : *t=a\_vy* ? 'T'as vu ?' Mais *ty=a\_vy* ? 'Tu as vu ?' est également bien formé. Les deux réalisations ne sont pas complètement interchangeables, elles ne relèvent pas du même registre et ne remplissent pas forcément les mêmes fonctions discursives : par exemple, attirer l'attention avec la première, poser avec insistance une question (éventuellement rhétorique) avec la seconde : *nʒ̃\_mε*, *ty=a\_vy sa* ? 'Non mais, tu as vu ça ?'. Bien entendu, l'infinie diversité des individus et des situations fait que poser des règles dans ce domaine n'est ni possible, ni souhaitable.

<sup>121</sup> Voir ci-dessous pour ce qu'il faut entendre par « neutre ».

La troisième personne est la plus riche en formes distinctes. La forme la plus courante de 3SG/PL sujet masculin est *i* devant consonne: *i=vjẽ* ‘il vient’, *i=vjen* ‘ils viennent’, *i=plø* ‘il pleut’ — c’était la prononciation réputée correcte aux dix-septième et dix-huitième siècles — mais *il* s’entend. Devant voyelle, *il* est le plus courant au singulier: *il=ariv* ‘il arrive’. En débit rapide, *i* est néanmoins possible, désyllabifié en la semi-voyelle [j]: *j=ariv u pa ?* ‘Il arrive ou pas?’<sup>122</sup> En revanche, le féminin court *ε* au lieu de *el* devant consonne — *ε=vjẽ* ‘elle vient’, *ε=vjen* ‘elles viennent’ — est stigmatisé à la différence de *i* qui, de fait, ne l’est pas, en dépit de condamnations puristes. A voir dans quelle mesure ce stigmate en restreint l’usage.<sup>123</sup>

Au pluriel masculin et féminin, devant voyelle, on entend /i(l)z|elzariv/. La question se pose à nouveau de l’identité morphologique du phonème /z/. Je propose d’y voir le même morphe proclitique identifié plus haut, d’où en transcription *i(l)=z=|el=z=ariv* ‘ils|elles arrivent’. De même que dans les suites Det-SN (*le=z=ami* ‘les amis’), l’insertion de *z* s’impose pour produire la signification plurielle — d’où peut-être la généralisation suivante: dans un groupe morphosyntaxique (SN ou SV) mettant en jeu un exposant clitique (déterminant ou clitique personnel), *z* est nécessaire à l’expression de la pluralité. Dans *le=z=ami*, il y a un double marquage, puisque *le* est déjà une forme plurielle. En revanche, on doit considérer les pronoms sujets de 3<sup>e</sup> personne comme sous-spécifiés pour le nombre, propriété évidente devant consonne où, *z* exclu, singulier et pluriel sont indistincts aux deux genres, entièrement si la forme verbale elle-même ne marque pas le nombre: *i(l)|el=fãt* ‘il(s)|elle(s) chante(nt)’. Cette analyse ne convient pas pour les pronoms 1PL et 2PL qui, malgré l’étiquette, ne sont pas réellement pluriels: il est bien connu que *nu* n’est pas le pluriel de *zə*, et *vu*, en tant que forme de politesse, peut être singulier. Je considère donc *nuz* et *vuz* comme des allomorphes, des formes longues de *nu* et *vu* devant voyelle.

Les formes dites neutres ont pour propriété sémantique de ne pas faire référence à des SN dénotant des entités individuelles, mais à des propositions généralement implicites ou bien à des ensembles. D’où la différence entre (a) *la=lẽgwistik, sa=fẽ\_pase lə=tã* ‘La linguistique, ça fait passer le temps’ et (b) *la=lẽgwistik, el=a\_fẽ boku\_də=progrẽ* ‘La linguistique, elle a fait beaucoup de progrès’. En (a), *sa* renvoie non à l’individu abstrait « la linguistique » comme le fait *el* en (b), mais à la proposition implicite « faire de la linguistique ». De même, *le=gato, sa=bur* ‘Les gateaux, ça bourre’ ou *le=gato, s=ε\_buratif* ‘Les gateaux, c’est bourratif’, où la référence de *le=gato* est générique, s’oppose à *le=gato, i(l)z=đ\_tro=kwi* ‘Les gateaux, ils ont trop cuit’, où il est question de gateaux individualisés. Dans *sa=sə=pure* ‘Ça se pourrait’, *sa* peut renvoyer à tout un discours. Dans *sa=va ?* ‘Ça va ?’, il signifie ce dont on doit s’enquérir d’une personne connue, que les participants à l’échange sont censés avoir présent à l’esprit, mais qu’aucune proposition close ne saurait exprimer.

Quoique non sans affinités, on ne confondra pas *sa* pronom personnel et *sa* pronom démonstratif (cf. 8.2). Les deux cooccurrent dans *sa, sa=mə=plẽ* ‘Ça, ça me plaît’. On aura noté que le pronom personnel ne figure que dans la case sujet du tableau 8.1. C’est que, partout ailleurs, on aurait affaire au démonstratif, toujours affecté d’une propriété déictique que le pronom n’a pas. Il est absent du tableau 8.2 pour la même raison. Quant à *sə* ‘ce’, il ne peut lui aussi que faire fonction de sujet, et seulement à la copule ÊTRE. Les exceptions relèvent de l’archaïsme: *syr=sə* ‘sur ce’, *sə\_mə=sãbl* ‘ce me semble’, le premier commun, le second affecté.

Bien qu’inscrits dans la case 3SG, *sə* et *sa* sont sous-spécifiés quant au nombre, aptes à renvoyer à des antécédents morphologiquement singuliers aussi bien que pluriels. Dans un registre influencé par le FO, cette sous-spécification a pour effet que *sə* peut-être suivi de la 3<sup>e</sup>

<sup>122</sup> *l=ariv* pour *il=ariv*, souvent cité dans les travaux sur le français « populaire », me semble obsolète. Mais cela reste à vérifier.

<sup>123</sup> La prononciation [a] réputée typique du « peuple » parisien n’est plus en usage, me semble-t-il — ou alors par plaisanterie: cf. *aboujphu* chez Queneau (*Zazie dans le métro*).

personne du pluriel de la copule si l'attribut est pluriel : *sə-sḍ de=lingwist* 'Ce sont des linguistes'. Cet accord inversé bénéficie de l'opprobre puriste qui pèse sur *s=ε de=lingwist* 'C'est des linguistes', pourtant parfaitement grammatical.<sup>124</sup> *sə* et *sa* sont également sous-spécifiés quant au genre grammatical, mais *sa* introduit une note affective, souvent péjorative, quand son antécédent dénote un ensemble d'humains : *le=zæn, sa=li=ply* 'Les jeunes, ça lit plus'.

En fonction de sujet des formes à initiale vocalique de la copule (présent et imparfait de l'indicatif), *sə* avec schwa toujours effacé est seul attesté dans le FP dont j'ai la compétence : *la=lēgwistik, s=ε | s=εte\_ēteresā* 'La linguistique, c'est|était intéressant'. Devant les formes de la copule à initiale consonantique, il y a variation entre *sə* et *sa* : *la=lēgwistik, sə= ~ sa=səre\_ēteresā si...* 'La linguistique, ce ~ ça serait intéressant si...'. L'emploi de *sə*, pour être bien implanté en FP, n'en est pas moins plus proche du FO oralisé, comme le montre l'anomalie de *sa* devant les formes indiscutablement propres à celui-ci : ? *sa=fy* au lieu de *sə=fy* 'ce fut'.

Devant les formes d'AVOIR auxiliaire, la variation est entre *s* et *sa*, sans qu'on sache bien si *s* est la forme courte de *sə* ou bien de *sa* : *la=lēgwistik, s= ~ sa=a\_ete\_ēteresā* 'La linguistique, ç'a ~ ça a été intéressant', *la=lēgwistik, s= ~ sa=ɔre\_ete\_ēteresā si...* 'La linguistique, ç'aurait ~ ça aurait été intéressant, si...'. On remarque en outre une contrainte (morpho)phonologique : /saete/ est sans problème, sans doute parce que facile à segmenter en ligne (*processing*) ; /saabjēmarʃe/ pour *sa=a\_bjē\_marʃe* 'Ça a bien marché', en revanche, semble préféré à ?/sabbjēmarʃe/, pour la même raison. Devant AVOIR verbe plein, non auxiliaire, *sa* paraît plus naturel : *la=lēgwistik, sa=a yn=lōg\_istwar* 'La linguistique, ça a une longue histoire' vs. ?*la=lēgwistik, s=a yn=lōg\_istwar* ? 'La linguistique, ç'a une longue histoire' — mais la différence est souvent difficilement audible. (On voit que dans ces exemples *el* 'elle' est également acceptable — cf. *la=lēgwistik, el=səre\_ēteresāt si...* 'La linguistique, elle serait intéressante, si...' — moyennant la différence sémantique indiquée ci-dessus, morphologiquement marquée par l'accord féminin de l'adjectif.)

La même variation *sə ~ sa* que devant la copule, avec pour *sə* la même proximité au FO oralisé (voire davantage), se manifeste devant les verbes modaux : *la=lēgwistik, sə= ~ sa=dwa(t) | pə(t)\_etr\_ēteresā* 'La linguistique, ce ~ ça doit | peut être intéressant'. Devant les autres types de verbes — ni copule, ni auxiliaires, ni modaux — seul *sa* est possible : *la=lēgwistik, sa=m=ēteres* 'La linguistique, ça m'intéresse'. On le trouve aussi avec les verbes météorologiques, en concurrence avec *i(l)* : *i=plø* 'il pleut' ~ *sa=plø* 'ça pleut', *i=nεz* 'il neige' ~ *sa=nεz* 'ça neige', etc. — où l'on voit que, contrairement à une idée reçue, *sa* n'est pas simplement l'équivalent « familier » de *səla* 'cela' qui, pour autant qu'il relève du FP, est exclu de ce contexte.

Pronom faible de 3<sup>e</sup> personne, *ḍ(n)* est singulier (collectif) quant au nombre : *ḍ=n=ε=pa\_serjə kã(t) ḍn=a\_diset\_ã* 'On n'est pas sérieux quand on a dix-sept ans'. Remplaçant *nu* 'nous', il est morphologiquement singulier du point de vue de l'accord verbal, mais sémantiquement sous-spécifié (collectif ou pluriel) quant à la concordance des prédicats adjectivaux variables en nombre : cf. *ḍn=ε\_zenjal ~ zenjo dā=la=famij* 'On est génial ~ géniaux dans la famille'.<sup>125</sup> L'expression synonyme *nu=səm\_zenjo...* 'Nous sommes géniaux', quoique sans doute moins fréquente et stylistiquement plus « marquée », n'en appartient pas moins au FP. *Nu=səm\_zenjal* 'Nous sommes génial', plutôt qu'un solécisme —

<sup>124</sup> Cf. « Oh ! C'est de singuliers pistolets ! » (Balzac, *Illusions perdues*, p. 227 du Livre de Poche).

<sup>125</sup> On nomme « accord » un processus obligatoire par lequel A et B syntaxiquement reliés expriment conjointement certaines propriétés indépendamment du sens. Cf. *le=sizo sḍ\_kase* 'Les ciseaux sont cassés' : le sujet pluriel, quoique dénotant un objet unique, contraint la pluralité verbale. La concordance, en revanche, se fonde sur le sens et elle est le plus souvent variable : cf. *s=ε mwa ki sḍi\_vəny* 'C'est moi qui suis venu' (concordance) ou *s=ε mwa ki ε\_vəny* 'C'est moi qui est venu' (accord avec le pronom relatif).



notion normative dont je n'ai pas l'usage : tout ce qui se dit est de la langue — procède de la généralisation de l'invariabilité des adjectifs (cf. chapitre 7).

**8.1.1.2. Les pronoms personnels faibles OD** — Ils suivent les pronoms sujets pour se procliticiser au verbe principal ou à l'auxiliaire :  $zə=l=ε$  *rākōtre ozurduji*,  $zə=la=rākōtrərə$  *ākər dəmə* 'Je l'ai rencontrée aujourd'hui, je la rencontrerai encore demain'. Comme on voit, la distinction du genre grammatical à la 3<sup>e</sup> personne du singulier est neutralisée devant voyelle. Transposer au pluriel nous donne  $zə=le=z=ε$  *rākōtre ozurduji*,  $zə=le=rākōtrərə$  *ākər dəmə* 'Je les ai rencontré.e.s aujourd'hui, je les rencontrerai encore demain'. (Cf. 8.1.1.6 pour le groupe clitique  $zə=lə||la|le$  ; et chapitre 11 pour la forme du participe passé *rākōtre*.) Le /z/ dans  $le=z=ε$  est le même exposant clitique vu ci-dessus avec les pronoms sujets. Avec un verbe impératif, les pronoms OD sont enclitiques. Ils se trouvent alors sous l'accent tonique qui affecte toute syllabe finale de groupe syntaxique. Cela a pour effet que les pronoms faibles 1SG et 2SG se changent alors en leurs contreparties fortes : *rəgard=mwa* ! 'Regarde-moi !', *rəgard=twa* ! 'Regarde-toi !', *twa* alors forcément réfléchi (cf. 8.1.1.4.) Les pronoms de 3<sup>e</sup> personne restent identiques à eux-mêmes : *Rəgard=lə||la|le* ! 'Regarde-le||la|les !', *lə* prononcé /lœ/ comme il se doit ; et de même les pronoms 1/2PL.

**8.1.1.3. Les pronoms personnels faibles OI** — Leurs propriétés clitiques sont les mêmes que celles des pronoms OD :  $zə=lɥi|lœr=ε$  *di kə*... 'Je lui|leur ai dit que...', *āvwa=lɥi|lœr\_ε=mel* 'Envoie-lui|leur un mail'. Quoique *lɥi* OI présente la même forme que *lɥi* fort, il est nécessaire de les distinguer, non seulement pour une raison syntaxique, mais aussi parce que le fort inclut une propriété de genre masculin, alors que le faible est épïcène. A noter que le clitique pluriel *z* peut s'insérer très normalement entre *lœr* et une voyelle :  $zə=lœr=z=ε$  *di* 'Je leur z-ai dit'. Cette forme est toutefois fortement stigmatisée — alors qu'elle est normale avec *lœr* déterminant possessif :  $lœr=z=ami$  'leurs amis' (voir chapitre 6).

**8.1.1.4. Les pronoms personnels faibles réfléchis** — Quant à la forme, il n'en existe qu'un, de 3<sup>e</sup> personne, sous-spécifié pour le nombre et le genre grammatical :  $ε|i(l)=sə=rəgard$  'Elle(s)|Il(s) se regarde(nt)'. Aux autres personnes, il n'existe pas de forme particulière distincte de la forme OD :  $ty=tə=rəgard$  'Tu te regardes'.

**8.1.1.5. Les pronoms personnels faibles obliques** — J'appelle « obliques » les pronoms compléments de prépositions. Seuls deux sont faibles : *i* 'y' et *ā* 'en' pour lesquels préposition et pronom se confondent :  $i = [dā/a + 3]$ ,  $ā = [də + 3]$ . En dépit des prescriptions de la grammaire académique, on les utilise couramment pour faire référence à des personnes : *frāswaz*,  $zə=m=ā=suvjē$   $ε$   $z=i=pās$  *suvā* 'Françoise, je m'en souviens et j'y pense souvent'. On peut hésiter quant à l'identité du /i/ qui apparaît dans les expressions stigmatisées mais fréquentes telles que  $z=i=ε$  *done*  $sō=kado$  'J'y ai donné son cadeau' : pronom oblique faible en fonction de pronom OI, ou bien réduction phonologique de *lɥi* 'lui' ?

**8.1.1.6. Les groupes clitiques** — Quand un pronom faible OD et un pronom faible OI cooccurrent, ils forment entre eux un groupe prosodique (ainsi qu'éventuellement avec un pronom faible sujet, toujours premier). Des contraintes strictes pèsent alors sur leur linéarisation réciproque selon la personne, ainsi que sur les groupes possibles.

Quant à ce dernier point, seuls sont admis les groupes {3 3} et {1|2 3}. Les groupes {1|2 1|2}, quoique sémantiquement concevables, sont exclus de quelque façon qu'on les ordonne : cf.  $*i(l)=mə=tə=kōfi$  '\*Il me te confie' ou  $*i(l)=tə=mə=kōfi$  '\*Il te me confie', seulement exprimable sous la forme  $i(l)=mə=kōfi$   $a=twa$  'Il me confie à toi' ou  $i(l)=tə=kōfi$   $a=mwa$  'Il te confie à moi'. Les groupes {1 1} et {2 2} le sont encore davantage, si possible : cf.  $*i(l)=nu=nu=kōfi$  '\*Il nous nous confie' pour dire  $i(l)=nu=kōfi$   $a=nu(=mem)$  'Il nous confie à nous(-mêmes)'. Tous ces énoncés déviants sont pourtant parfaitement compréhensibles. La restriction est bien morphologique, elle ne tient pas au sens.

Dans les groupes {3 3}, OD précède OI, toujours si le groupe est proclitique :  $i(l)=lə=lɥi=kōfi$  'Il le lui confie', variablement s'il est enclitique (à l'impératif) :

*kɔ̃fi=lə=lyi ! ~ kɔ̃fi= lyi=lə !* ‘Confie-le lui’ ~ ‘Confie-lui le’. L’ordre inverse (OI < OD) prévaut dans les groupes {1|2 3} : *i(l)=mə=lə=kɔ̃fi* ‘Il me le confie’, avec la même variation à l’impératif : *kɔ̃fi=lə=mwa ! ~ kɔ̃fi=mwa=lə !* ‘Confie-le moi !’ ~ ‘Confie-moi le !’.

Les groupes proclitiques se laissent accroître d’un pronom faible 2SG dit « éthique », dont la fonction est de prendre l’auditoire à témoin : *ʒə=vɛ tə=lə=lyi=rãdr (sɔ̃=fify\_bukê)* ‘Je vais te le lui rendre (son fichu bouquin)’. Dans *ty=va mə=la=lave, set=vesel* ‘Tu vas me la laver, cette vaisselle’ ou *ty=va tə=la=rãze, ta=fãbr* ‘Tu vas te la ranger, ta chambre’, le pronom 1 ou 2 remplit aussi une fonction « applicative » (« faire pour... ») qui, dans de nombreuses langues, s’exprimerait au moyen d’une forme spéciale du verbe.

Le groupe *lə|la=lyi* se réduit souvent à /lyi/, voire /i/ : cf. *ʒə=lə=lyi=rãdre* ‘je le lui rendrai’ prononçable /ʒɛlɛlyiãdrɛ/ ~ /ʒlɛlyiãdrɛ/ ~ /ʒɛllɛlyiãdrɛ/ ~ /ʒiãdrɛ/. On peut voir là un phénomène phonologique lié au débit. Mais il peut aussi s’agir d’un processus morphosyntaxique, dans la mesure où ne sont mis en jeu que des verbes à double objet dénotant une translation (*ãvwaje* ‘envoyer’, *done* ‘donner’, *rãdr* ‘rendre’, etc.), dont le sens fait du destinataire le complément principal. Il se peut donc que, plutôt que *ʒə=lə=lyi=rãdre*, la forme non réduite de /ʒiãdrɛ/ soit en fait *ʒə=lyi=rãdre* ‘je lui rendrai’, où seul le destinataire est exprimé, l’objet direct (patient, thème) restant implicite.

**8.1.1.7. Les pronoms personnels forts** — Ils remplissent trois fonctions : (a) thématique (*mwa, ʒə...* ‘Moi, je...’) ; (b) prédicative (complément de la copule, cf. *s=ɛ mwa* ‘C’est moi’) ; (c) compléments de prépositions (obliques) (*avɛk mwa* ‘avec moi’). Quant aux deux premières, la morphologie n’a rien à en dire. Pour la troisième, on remarque l’inexistence de contreparties fortes de *i* ‘y’ et *ã* ‘en’ (voir plus haut). Et aussi que le réfléchi *swa* ‘soi’ passe mal en fonction de thème, ou alors sous la forme amplifiée *swa=mɛm* ‘soi-même’ : comparez *ɔ̃=pãs a=swa(=mɛm) avã də=pãse o=z=otr* ‘On pense à soi(-même) avant de penser aux autres’ et *swa=mɛm, ɔ̃n=i=pãsarɛ=pa* ‘Soi-même, on n’y penserait pas’, qui sonne mieux à mon oreille que *swa, ɔ̃n=i=pãsarɛ=pa*. En fonction prédicative, on dit facilement *i(l)=fo(t) ɛtr swa(=mɛm)* ‘Il faut être soi(-même)’. Dans les contextes autres que génériques il est toujours possible de substituer *lyi* ‘lui’ ou *ɛl* ‘elle’ à *swa* : *sofi (nə)=pãs\_k=a=swa ~ ɛl(=mɛm)* ‘Sophie ne pense qu’à soi ~ elle(-même)’.

## 8.2. Les pronoms démonstratifs

Cinq formes : M.SG *səlyi ~ syi* ‘celui’, M.PL *sə* ‘ceux’, F.SG/PL *sɛl* ‘celle(s)’, NTR *sə* ‘ce’ et *sa* ‘ça’. Pronom démonstratif — à distinguer du pronom personnel et du déterminant démonstratif —, le neutre *sə* est limité à la fonction d’antécédent d’une proposition relative : *sə\_kə ʒə=vɔ...* ‘ce que je veux’, *sə\_ki mə=trakas...* ‘ce qui me tracasse...’, fonction d’où *sa* est exclu en FP « hexagonal », à moins qu’il ne fasse lui-même fonction de prédicat : *s=ɛ sa ki mə=ple* ‘C’est ça qui me plaît’. Quant aux suites *sə\_kə* et *sə\_ki*, on peut se demander si elles n’ont pas été réanalysées en des formes simples *skə* et *ski* analogues à l’anglais *what* (cf. *what I want..., what nags me...*).

Une telle réanalyse paraît probable pour la particule interrogative *ɛskə... ?* ‘est-ce que... ?’, considérant que les formes fléchies « était-ce que... ? », « sera-ce que... ? » sont quasi obsolètes même en FO oralisé.<sup>126</sup> (Seul « serait-ce que... ? » semble encore usuel, quoique affecté.) A quoi s’ajoute que l’interrogation inversée du type de FO « Est-ce lui ? », « Est-ce vrai que... ? » est pour ainsi dire inusitée en FP spontané. Reste que la perméabilité du FP au FO (écrit et oralisé) fait que ces tournures lui appartiennent dans une certaine

<sup>126</sup> « La locution *est-ce que...* Elle est très ancienne dans la langue ; elle avait à l’origine une valeur sémantique fortement marquée, « est-il vrai que », « est-il possible que », elle est devenue une lexie inanalysable dans le détail de ses éléments, un outil grammatical spécifique de l’interrogation directe » (Spillebout 1985 : 381). La réanalyse est achevée dans la plupart des créoles de base lexicale française : cf. haïtien *Eske ou vle... ?* ‘Est-ce que tu veux...?’.

mesure, et maintiennent peut-être une connaissance subliminale du caractère construit, plurimorphique de la séquence. Et voir aussi l’alternance entre /kɛs(kə)tyvø/ ‘Qu’est-ce que tu veux ?’ — dont l’analyse délivrée de l’orthographe ne va pas de soi, j’y reviendrai en 8.4 — et *s=ɛ kwa kə ty=vø* ? ‘C’est quoi que tu veux ?’, à la morphosyntaxe plus transparente.

Quant au trois formes non neutres, on ne cèdera pas à la tentation d’y voir la composition du déictique *sə* / *sɛ(t)* avec les pronoms forts *lyi*, *ø* et *ɛl*, « analyse » cognitivement peu plausible en synchronie et diachroniquement inexacte (cf. Raynaud de Lage 1967 : 47-48). Les formes simples ne peuvent servir que d’antécédent d’une proposition relative (*sɛl ki...* ‘celle qui...’). Sujets ou objets d’un verbe ou compléments d’une préposition, elles doivent être amplifiées d’un élément déictique enclitique =*si* ‘ci’ ou =*la* ‘là’ selon l’éloignement physique ou mental du référent: cf. *nu=prād-rḡ səlyi=si*, *pa səlyi=la* ‘Nous prendrons celui-ci, pas celui-là’.

Ces mêmes extensions permettent au pronom démonstratif neutre *sə* de faire fonction d’argument : cf. *nu=prād-rḡ səsi*, *pa səla* ‘Nous prendrons ceci, pas cela’. Je transcris à l’instar de l’orthographe *səsi* et *səla* comme un seul mot, car il est probable que leur complexité n’est plus sentie. Il semble que *səsi* soit d’usage plus rare que son équivalent plus « individualisant » *səlyi=si*, et que la distinction en termes de distance soit souvent neutralisée par l’emploi unique de *sa* : *nu=prād-rḡ sa*, *pa sa* ‘Nous prendrons ça, pas ça’ (geste).

### 8.3. Les pronoms relatifs

Les deux pronoms relatifs fondamentaux du FP sont *ki* ‘qui’ et *kə* ‘que’. Ils ne se fléchissent ni pour le genre ni pour le nombre. Quant au cas, *ki* est sujet ou bien oblique, régi par une préposition. Sujet, il accepte tout type d’antécédent : *lə=fə ki mjol dā=la=kur* ‘le chat qui miaule dans la cour’, *la=fɛz ki s=ɛ kase* ‘la chaise qui s’est cassé’ ; oblique, son antécédent doit dénoter une entité animée : cf. *lə=fə a=ki gəstḡ aʃɛt dy=mjaumjam* ‘Le chat à qui Gaston achète du Miaou-Miam’ vs. *\*la=fɛz syr=ki zə=m=aswa* \*‘la chaise sur qui je m’assoie’. Mais encore faut-il que ledit animé soit placé assez haut sur l’échelle commune (culturelle, non scientifique) du vivant : cf. *#la=furmi syr=ki z=ɛ marʃɛ* ‘#la fourmi sur qui j’ai marché’ — sauf personnalisation toujours possible : cf. *la=sigal a=ki la=furmi rəʃyz sən=ɛd* ‘la fourmi à qui la cigale refuse son aide’.

Plutôt qu’à deux formes du même lexème QUI, on peut donc estimer avoir affaire à deux lexèmes distincts QUI1 et QUI2, QUI1 pronom relatif sujet non spécifié pour le trait d’animéité ([CAT *pro*, REL +, CAS *sujet*]), QUI2 pronom relatif oblique spécifié positivement pour ce trait ([CAT *pro*, REL +, CAS *obl*, ANIM +]). Une autre possibilité consiste à mettre en œuvre une contrainte sur la co-occurrence des propriétés : [CAT *pro*, REL +, CAS *obl*] ⇒ [ANIM +].<sup>127</sup> Qu’on opte pour l’une ou l’autre solution, on doit être conscient qu’on ne fait que traduire en langage formel ce qui se décrit aussi bien, quoique plus prolixement, en langage courant. L’explication est à chercher dans l’histoire de la langue. Le moyen français était moins strict, qui tolérait « le doux traict par qui je fus blessé » (Gougenheim 1951 : 91). En fonction de sujet, *qui* et *que* alternaient, variantes dialectales plus que formes distinctes : cf. « ... plusieurs choses que, pour l’eure, ne seront pas racontées » (Zink 1990 : 47). Hormis proverbes et dictons — p.ex. *ki s=i=frɔt s=i=pik* ‘Qui s’y frotte s’y pique’ — l’emploi sans antécédent de *ki* ‘qui’ n’appartient pas ou peu au FP.

Le pronom relatif *kə*, en revanche, ne peut être régi par une préposition. Lorsque le caractère inanimé (ou insuffisamment animé du point de vue de l’échelle commune) de l’antécédent exclut *ki*, c’est *kwa* ‘quoi’, forme forte analogue à *twa* vis-à-vis de *ty*, qui le remplace : cf. *la=fɛz syr=kwa zə=m=aswa* ‘la chaise sur quoi je m’assoie’ ;<sup>128</sup> ou bien, plus

<sup>127</sup> *Feature Cooccurrence Restriction* (FCR) dans le modèle syntaxique *Generalized Phrase Structure Grammar* (GPSG) (Gazdar et al. 1985 ; Abeillé 1993).

<sup>128</sup> Le moyen français acceptait « la dame de quoy nous parlons » (Zink 1990 : 47).

proche du FO oralisé, un membre de la série *ləkəl / lakəl / lekəl* ‘lequel / laquelle / lesquels / lesquelles’, formant mot-valise avec les prépositions *a* ‘à’ et *də* ‘de’ : *la=fɛz syr=lakəl zə=m=aswa* ‘la chaise sur laquelle je m’assoie’, *S=et\_a=l=amur okəl zə=pās* ‘C’est à l’amour auquel je pense’ (Françoise Hardy, 1962) ;<sup>129</sup> ou bien encore (esquivant le problème) *la=fɛz u zə=m=aswa* ‘la chaise où je m’assoie’. Mais la solution la plus typique du FP est d’employer *kə* dont on doit considérer qu’il ne porte d’autre trait casuel que négatif, [CAS –*sujet*], ce qui le rend disponible pour toutes les autres fonctions (voir ci-dessous).

L’emploi de *kə* en fonction d’OD ne différencie pas le FP du FO : cf. *lə=fə kə gəstō nuri o=mjaumjam* ‘Le chat que Gaston nourrit au Miaou-Miam’. En fonction d’OI datif mettant en jeu la préposition *a* ‘à’, *kə* se voit associé dans la proposition relative à un pronom dit « résomptif », lié à l’antécédent et exprimant le trait casuel : cf. *lə=fə kə gəstō lɥi=afɛt dy=mjaumjam* ‘Le chat que Gaston lui achète du Miaou-Miam’, qu’on comparera à la construction plus normative donnée plus haut. Quoique stigmatisée, cette construction, du reste bien représentée dans les langues du monde, est assez commune en parole peu surveillée. La langue du seizième siècle l’acceptait sans sourciller (cf. Zink 1990 : 47-48).

En fonction d’oblique, deux cas sont à distinguer selon la nature de la préposition. S’il s’agit de *də*, le FP peut, comme le FO, avoir recours au pronom relatif oblique *dō* ‘dont’, remplaçable par *də=ki* ‘de qui’ si les conditions décrites ci-dessus s’y prêtent : cf. *lə=rɔmā dō z=ɛ\_dɛfɪrɛ la=kuvɛrtɪr* ‘le roman dont j’ai déchiré la couverture’, *lə=rɔmā dō tuləmōd parl* ‘le roman dont tout le monde parle’, *l=otær dō ~ də=ki tuləmōd parl* ‘l’auteur dont ~ de qui tout le monde parle’, *le=dokymā dō z=ɛ\_bəzwɛ* ‘les documents dont j’ai besoin’. Également disponible, la stratégie résomptive est propre au FP : cf. *lə=rɔmā kə z=ɛ\_dɛfɪrɛ sa/la=kuvɛrtɪr* ‘le roman que j’ai déchiré sa/la couverture’, *lə=rɔmā kə tuləmōd ā parl* ‘le roman que tout le monde en parle’, *l=otær kə tuləmōd ā parl* ‘l’auteur que tout le monde en parle’, *le=dokymā kə z=ɛ\_bəzwɛ* ‘les documents que j’ai besoin’. On voit que tout élément de reprise peut être absent, facultativement — et de fait souvent — dans le premier exemple (*la* au lieu de *sa*), presque toujours dans le dernier, peut-être parce que *də* n’y a aucun contenu sémantique propre, partie indispensable de la locution lexicalisée *avwar bəzwɛ də* ‘avoir besoin de’. Je ne suis pas tout à fait sûr, en revanche, que le deuxième et le troisième exemple soient possibles sans *ā*. Une étude de corpus oral devrait permettre de le vérifier. Il est probable qu’on les y trouvera attestés.

Avec une autre préposition hormis *a* ‘à’, *kə* s’associe à la variante intransitive de celle-ci : *la=fɛz kə zə=mə=sɥi\_asi(z) dəsy* ‘la chaise que je me suis assis.e dessus’. Mais ces tournures sont particulièrement condamnées par la norme. On les contourne aisément : cf. *la=fɛz u zə=mə=sɥi\_asi(z)* ‘la chaise où je me suis assis.e’, qui permet en outre d’éviter les plus complexes *syr=kwa* ‘sur quoi’ et *syr=lakəl* ‘sur laquelle’.

#### 8.4. Les pronoms interrogatifs

Je ne considérerai que les trois pronoms *ki* ? ‘qui ?’, *kə* ? ‘que ?’ et *kwa* ? ‘quoi ?’. Leur forme et des considérations étymologiques font qu’on les range souvent avec les pronoms relatifs sous l’étiquette commune de « mots *qu-* », transposition de l’anglais *wh-words* (cf. *who* ‘qui’, *what* ‘que, quoi’). L’homophonie ou, à tout le moins, la parenté étymologique des pronoms relatifs et interrogatifs est toutefois loin de constituer un trait général. En wolof (Niger-Congo, atlantique, parlé au Sénégal et en Gambie), par exemple, l’interrogatif est *-an* préfixée d’une marque de classe nominale : *k-an* ? ‘qui ?’, *l-an* ? ‘quoi ?’ : *K-an dem na* ? {CLK-INT partir ACP.3SG} ‘Qui est parti ?’ ; tandis que le relatif est *-u/i* accordé en nombre et

<sup>129</sup> Il y eut des puristes pour dénoncer le solécisme, au prétexte qu’on devrait dire (en FO oralisé) « C’est à l’amour que je pense » ou bien « C’est l’amour auquel je pense », en ne marquant qu’une fois le cas oblique. Ils ou elles auraient dû savoir que l’économie n’est pas une vertu des langues parlées. Boileau n’y prenait pas garde : « C’est à vous, mon esprit, à qui je veux parler » (Spillebout 1985 : 166).

en classe avec l'antécédent : *jigéen j-u dem j-i* {femme CLj-REL partir CLj-DEF} 'la femme qui est partie'.

Seuls *kə ?* et *kwa ?* posent un problème aux morphologues du fait qu'ils ont même contenu sémantique schématisable par l'ensemble de traits [Q + HUM -], tandis que *ki ?* se laisse caractériser comme [Q + HUM +], moyennant les mêmes vacillations que pour le relatif *ki* : cf. Q *ki ε dāvā la-pōrt ?* 'Qui est devant la porte ?' — R #*lə=fə* 'Le chat' vs. *ki ε dāvā la-pōrt ?* 'Qui est devant la porte ?' — R \**?yn=suri* 'Une souris'. Le fait est qu'il n'existe pas en français (ni dans bien des langues) d'interrogatif réellement adéquat pour renvoyer à des animés non humains. Quant au cas, *ki ?* et *kwa ?* sont non-spécifiés : cf. *ki | kwa ε tōb-e ?* 'Qui | Quoi est tombé ?' (le second plutôt question-écho avec prosodie idoine), *t(y)=a vy ki | kwa ?* 'Tu as vu qui|quoi ?', *a=ki | kwa ty=pās ?* 'A qui|quoi tu penses ?', etc. Il est de tradition de considérer *kə* comme la forme faible de *kwa* de la même façon que *tə* 'te' l'est de *twa* 'toi'.

Pour *kə ?* deux analyses paraissent possibles, l'une qu'on peut dire « FO-compatible », l'autre qui creuse l'écart entre FO et FP. Tout d'abord, en amont de l'alternative, il est indubitable que les deux systèmes partagent une forme *kə*, dont le schwa s'élide devant voyelle et qui présente deux propriétés : (a) être limitée à la fonction grammaticale OD : cf. \**Kə ε tōb-e ?* \*'Que est tombé ?' ; (b) n'être admissible qu'en tête d'une proposition dont le verbe (l'auxiliaire si périphrase) occupe la deuxième position, d'où « inversion » du sujet : cf. *k a t ε l di ?* 'Qu'a-t-elle dit ?' vs. \**ε l = a di kə ?* \*'Elle a dit que ?', *kə di Marsəlin ?* 'Que dit Marceline ?' vs. \**kə Marsəlin di ?* \*'Que Marceline dit ?'. La seconde propriété s'accorde en partie avec l'analyse reçue mentionnée ci-dessus. On explique ainsi l'agrammaticalité de \**ε l = a di kə ?* et de \**a = kə ty = pās ?* \*'A que tu penses ?', ni la position OD canonique ni celle de complément d'une préposition n'étant des positions faibles. On peut ainsi caractériser *kə ?* par l'ensemble de traits suivants : [CAT *pro*, Q +, FORME *faible*, HUM -, INV +, CAS *od*, EXTR +], avec valeur positive du trait [extr(action)].

L'autre analyse vaut de la variété qui ne pratique pas l'inversion simple (*k = a di marsəlin ?* 'Qu'a dit Marceline ?') et ne produit que *marsəlin a di kwa ?* 'Marceline a dit quoi ?' (Q dit « in situ »), /kɛskə/ *marsəlin a di ?* 'Qu'est-ce que Marceline a dit ?' ou /kɛsk/ *a di marsəlin ?* 'Qu'est-ce qu'a dit Marceline?'.<sup>130</sup> Là, *kə* est remplacé par /kɛskə/ ou /kɛski/ si l'objet de la question est non-humain et de fonction OD ou sujet : cf. /kɛski/ *ε tōbe ?* 'Qu'est-ce qui est tombé ?' ; et de même, mais moins régulièrement, /kiɛskə/ ou /kiɛski/ au lieu de *ki* si l'objet de la question est humain et OD ou sujet : /kiɛskə/ *t(y)=a vy ?* 'Qui est-ce que tu as vu ?', /k(i)jɛski/ *ε tōbe ?* 'Qui est-ce qui est tombé ?'.<sup>131</sup> Les transcriptions phonologiques sont toutes provisoires, elles entendent ne pas préjuger de l'analyse de ces formes. Devant un pronom faible à initiale consonantique, /kɛskə/ et /kiɛskə/ peuvent s'abréger en /kɛs/ et /kiɛs/ et se cliticiser au pronom : /kɛs/ = *ty = fε ?* 'Qu'est-ce tu fais ?', /kiɛs/ = *ty = vwa ?* 'Qui est-ce tu vois ?'. Devant voyelle le /i/ final de /kɛski/ et /kiɛski/ peut tomber : /kɛsk/ = *ε tōbe ?* 'Qu'est-ce qu'est tombé ?', /kiɛsk/ = *ε tōbe ?* 'Qui est-ce qu'est tombé ?'. On voit mal comment expliquer ces phénomènes par des processus phonologiques actifs en synchronie, même s'il est indéniable qu'ils sont aussi liés au débit. (Même en débit rapide, une variété plus alignée sur le FO oralisé reculerait devant /kɛs/ et n'irait pas non plus au-delà de /kɛskjɛtōbe/ par désyllabication du /i/.)

Cela étant, deux analyses s'offrent à nous. La plus radicale consiste à considérer ces quatre formes comme autant de pronoms interrogatifs synchroniquement inanalysables, mais

<sup>130</sup> Entendons-nous bien : cette variété existe, mais seulement en tant qu'abstraction analytique. Je veux dire par là qu'il est hautement improbable qu'aucun francophone n'ait qu'elle dans sa compétence active et/ou passive.

<sup>131</sup> Trois prononciations s'entendent, selon le débit : /kiɛskə/i/, /kijɛskə/i/ et /kjɛskə/i/. En français classique *qui ?* s'employait au sens de *qu'est-ce qui ?* : p.ex. « Qu'avez-vous ? Qui vous fait soupirer ? » (Corneille, *Le Menteur*) (Spillebout 1985 : 172). C'est pourquoi le vers fameux de La Fontaine « Qui te rend si hardi de troubler mon breuvage ? » est généralement mal compris : le loup ne demande pas « qui ? », mais « quoi ? ».

fléchis pour le trait [ $\pm$ humain] et le cas. D'où le paradigme ci-dessous où l'on voit l'alternance /kiɛ/ | /kɛ/ associée à l'opposition d'homininité et l'alternance de la voyelle finale à la différence casuelle :

	+H	-H
suje t	kiɛski	kɛski
OD	kiɛskə	kɛskə

L'idée est séduisante, mais elle se heurte à un problème. Le tableau devrait en effet comporter une quatrième ligne pour la fonction oblique, soit pour les cas où ces « pronoms » complètent une préposition. On observe alors un intéressant contraste :  $a=/kiɛskə/ ty=pãs ?$  'A qui est-ce que tu penses ?' est bien formé, tout comme  $a=ki ty=pãs ?$  'A qui tu penses ?', alors que  $*a=/kɛskə/ ty=pãs ?$  '\*A qu'est-ce que tu penses ?' n'est pas une alternative possible à  $a=kwa ty=pãs ?$  'A quoi tu penses ?'. On en conclut que /kɛskə/ inclut kə et que l'agrammaticalité de  $*a=/kɛskə/$  répond à celle de  $*a=kə$  'à que'. Dans ces conditions, on se prend à douter que /kɛskə/ et les autres formes du paradigme soient indécomposables et l'on doit se poser la question de la nature de l'élément /ɛskə/i/ qui suit /k/ et /ki/.

Y répondre ne se peut sans prendre en compte ceci : la représentation orthographique donne à penser qu'on a affaire à une construction clivée interrogative — cf. « Est-ce James Joyce qui a écrit l'Odyssée ? » — remplissant la fonction discursive de focalisation typique des clivées. Ce fut sans doute vrai (cf. Gougenheim 1951 : 212). Mais, quant au présent, le fait est que /kiɛski/  $a\_ekri l=odise ?$  ne diffère guère de  $ki a\_ekri l'Odyssée ?$  Les deux phrases sont de simples demandes d'information à propos de l'identité de l'écrivain, sans y ajouter cette connotation d'insistance et d'incertitude — « Est-ce (vraiment) X (et non pas Y) qui... » — caractéristique des clivées interrogatives. En FP spontané les formes les plus couramment utilisées pour introduire cette connotation seraient plutôt  $s=\epsilon ki ki a\_ekri l=odise ?$  'C'est qui qui a écrit l'Odyssée ?' ou  $ki s=\epsilon ki a\_ekri l=odise ?$  'Qui c'est qui a écrit l'Odyssée ?'.

Peut-être convient-il aussi de noter que le segment /ɛs/ entre /ki/ initial et /ki/ final est susceptible de s'effacer : cf.  $Ki ki a\_ekri l=odise ?$  'Qui qui a écrit l'Odyssée ?' ? Il est vrai que cette tournure, à cause peut-être du kakemphaton [kiki], s'est trouvée en première ligne sous le feu du stigmatisme et n'est plus guère courante dans l'hexagone — elle le serait davantage au Québec.<sup>132</sup> Son existence n'en dévoile pas moins une possible individualité de /ɛs/. Mais elle en révèle aussi l'inanité sémantique dès lors que /kiki/ et /kiɛski/ apparaissent parfaitement équivalents, tandis que « Est-ce James Joyce qui a écrit l'Odyssée ? » et « James Joyce a écrit l'Odyssée ? » (avec intonation interrogative) ne le sont évidemment pas.

Et si, vu l'absence de focalisation, l'élément /ɛskə/ n'était autre que la particule interrogative  $\epsilon skə$ , étymologiquement complexe, mais que j'ai supposée plus haut être à présent compacte en FP ? Particule redondante dans ce contexte, mais cela ne saurait nous arrêter, tant redondance et langue orale font bon ménage. Mais que dire alors des finales /ki/ et /kə/ associées à la fonction grammaticale de l'objet de la question ? La particule n'est pas soumise à flexion. Reste à supposer qu'elle nous montre ici sa forme abrégée — cf.  $k/\epsilon s/=ty=f\epsilon ?$  'Qu'est-ce tu fais ?' — suivie des pronoms relatifs  $ki$  'qui' et  $kə$  'que', supposition confortée par la possibilité, assez limitée, d'insérer un adverbe ou une incise : cf.  $k/\epsilon s/_d\check{s}k\_kə=ty=f\epsilon ?$  'Qu'est-ce donc que tu fais ?',  $k/\epsilon s/_di\_ty\_k=il=a\_f\epsilon ?$  'Qu'est-ce, dis-tu, qu'il a fait ?', tournures peu idiomatiques, me semble-t-il, en FP spontané.<sup>133</sup>

<sup>132</sup> De même les questions du type de  $u kə ty=va ?$  'Où que tu vas ?',  $purkwa kə ty=di sa ?$  'Pourquoi que tu dis ça ?', etc., très stigmatisées et peu usitées, semble-t-il, de ce côté-ci de l'Atlantique.

Au final on en reviendrait à la structure notée par le FO, sauf pour la compacité de *es*, déduite de la quasi absence en FP des formes temporellement fléchies que le FO conserve — mais même là l'éventail flexionnel est limité : présent (<Qui/Qu'est-ce que/qui... ?>) et imparfait (<Qui/Qu'était-ce que/qui... ?>), à l'exclusion commune ou totale du futur (? <Qui/Que sera-ce que/qui... ?>), du passé simple (\*<Qui/Que fut-ce que/qui... ?>) et des temps composés (\*<Qui/Qu'a-ce été que/qui... ?>).<sup>134</sup> Il faudrait donc noter *Kə/Ki\_εs\_kə/ki... ?*, le schwa de *kə* toujours effacé devant l'initiale vocalique de *εs*.

Il reste toutefois un détail à ne pas négliger, qui ressort de la comparaison de *k\_εs\_kə/ki... ?* avec *kə\_di\_marsəlin ?* 'Que dit Marceline ?' et *\*kə Marsəlin di ?* 'Que Marceline dit ?' (voir ci-dessus) : la position initiale de *kə* se soutient du fait d'être, au moins historiquement, suivi immédiatement d'un prédicat à sujet inversé (FO « est-ce ») dont *kə* est le complément extrait. On remarquera à ce propos que, si l'adverbe ou l'incise est insérée avant *εs* plutôt qu'après, l'acceptabilité de *kə* se dégrade sensiblement : cf. *?kə, di=ty, [εs]k=il=a\_fε ?* 'Que, dis-tu, est-ce qu'il a fait ?'.<sup>135</sup> Pourtant les arguments en faveur de la présente indécomposabilité de *εs* sont trop solides pour qu'on souhaite y renoncer.

Si l'on n'y renonce pas, il reste un recours, à savoir s'en tenir au paradigme ci-dessus et analyser les propriétés morphosyntaxiques de ses formes comme résultant de la combinaison d'un sous-ensemble des propriétés de leurs composantes historiquement amalgamées. Ce qui laisse le choix quant à la représentation la plus adéquate : ou bien considérer que l'amalgame est total et que *kεskə*, etc. sont dépourvus de structure interne ; ou bien y voir des mots complexes, suites de morphes cliticisés les uns aux autres et inséparables les uns des autres en FP spontané. Le caractère quand même relatif de cette inséparabilité fait que je me range à cette dernière option, d'où le paradigme suivant enrichi d'une ligne pour la fonction oblique :

	+H	-H
sujet	ki=εs=ki	k=εs=ki
OD	ki=εs=kə	k=εs=kə
oblique e	P=ki=εs=kə	

La case vide du tableau est due, on s'en souvient, à la faiblesse de *kə* qui lui interdit d'être régi par une préposition, contexte où le remplace la forme forte *kwa*. Les cinq formes se laissent décrire par les ensembles de traits suivants :

*ki=εs=ki* : [CAT *pro*, Q +, FORME *ns*, HUM +, CAS *sujet*, EXTR +]

*ki=εs=kə* : [CAT *pro*, Q +, FORME *ns*, HUM +, CAS *od*, EXTR +]

*k=εs=ki* : [CAT *pro*, Q +, FORME *ns*, HUM -, INV -, CAS *sujet*, EXTR +]

*k=εs=kə* : [CAT *pro*, Q +, FORME *ns*, HUM -, INV -, CAS *od*, EXTR +]

*P=ki=εs=kə* : [CAT *pro*, Q +, FORME *ns*, HUM +, INV -, CAS *obl*, EXTR +]

*kə* étant faible, donc [EXTR -] et *ki* non spécifié pour la propriété [EXTR] — car initial ou *in situ* — ces pronoms héritent leur propriété [EXTR +] de la particule interrogative *εs(kə)*, pour autant que c'est une propriété caractéristique de ces particules que de se situer à la marge

<sup>133</sup> A ne pas confondre, évidemment, avec *kə di=ty k\_il=a\_fε ?* 'Que dis-tu qu'il a fait ?', où *kə* complète le second verbe, non le premier.

<sup>134</sup> « Sera-ce (que)... ? », « Serait-ce (que)... ? », « Fût-ce... » appartiennent bien sûr à un FO littéraire, voire académique, mais il s'agit d'une toute autre construction. Il en allait autrement en moyen français : cf. « Jusques à quand sera-ce que tu abuseras de la jeunesse de nostre Roy » (Gougenheim 1951 : 213).

<sup>135</sup> *kwa, di=ty, /εs/k=il=a\_fε ?* 'Quoi, dis-tu, est-ce qu'il a fait ?' vaut mieux, mais sonne toujours un peu gauche, du moins à mon oreille.

initiale ou finale des propositions sur lesquelles elles portent, du fait précisément de les avoir toutes entières dans leur portée (cf. Creissels 2006b : 170-173).<sup>136</sup> Ils héritent la valeur positive de la propriété [Q] (question) et la valeur plus ou moins de la propriété d'homininité des pronoms interrogatifs basiques *ki* et *kə*. La valeur de la propriété de cas leur vient des pronoms relatifs amalgamés *ki* et *kə* ou de la préposition (P). Enfin, ils sont non-spécifiés (ns) pour la propriété de forme fort ou faible : faibles, ils ne pourraient être initiaux ; forts, ils pourraient l'être, mais aussi *in situ*, ce qui, on l'a vu, est impossible. La propriété [INV] (inversion) n'est pertinente que pour les formes non-sujet ; elle est toujours négative : cf. *\*k=εs=k\_a=ty\_di ?* 'Qu'est-ce qu'as-tu dit ?'.

On a vu au passage que la particule interrogative est susceptible d'accompagner tout mot ou syntagme interrogatif extrait : *a=kwa (eskə) ty=pās ?* 'A quoi (est-ce que) tu penses ?', *kəl\_film (eskə) ty=prefer ?* 'Quel film (est-ce que) tu préfères ?', *u (eskə) ty=va ?* 'Où (est-ce que) tu vas ?', etc. Elle est exclue, en revanche, quand le pronom ou syntagme occupe sa position argumentale canonique : *\*eskə ty=prefer kəl\_film ?* '\*Est-ce que tu préfères quel film ?', *\*eskə ty=va u ?* 'Est-ce que tu vas où ?', etc. Autrement dit, *eskə* est ou bien isolé et initial dans une question totale — (*eskə) ty=va partir ?* '(Est-ce que) tu vas partir ?') — ou bien, si présent dans une question partielle, obligatoirement postposé à un mot ou un syntagme interrogatif extraits. Je laisse aux syntacticiens d'aller plus loin dans l'étude de ce phénomène.<sup>137</sup>

## 9. La flexion verbale : vue d'ensemble

A la différence des noms dont la flexion est semi-analytique ou supplétive, les verbes connaissent en outre une flexion synthétique fondée sur la combinaison d'une base et d'exposants suffixés : p.ex. *lav-e* 'laver | lavé.e(s)', où /lav/ est la base, /-e/ le suffixe infinitif ou participial ; et aussi une flexion analytique périphrastique ou « fusionnée » (Abeillé & Godard 2021, chapitre 3), par exemple *ɜ=ε\_lav-e* 'j'ai lavé'. On se reportera au chapitre 5 pour ces notions. Le choix du terme « base » plutôt que « racine » ou « radical » ou « thème » se justifie du fait que le FP n'offre guère de critères évidents permettant de distinguer ces trois unités. « Base » les recouvre de façon commode sans distordre la réalité. En outre, ce qui rend la flexion du FP particulièrement complexe et opaque, ce n'est pas la richesse des exposants, plutôt moyenne, mais bien la diversité souvent imprévisible des bases auxquelles s'attachent tels ou tels exposants flexionnels de temps-aspect-mode (TAM) et personne-nombre-genre (PNG).

Au plus simple et selon le critère de la diversité des bases, les verbes se répartissent entre deux classes : les verbes à une seule base (classe I) et les verbes à plusieurs bases (classe II).<sup>138</sup> *Lav-e* est un exemple de verbe de classe I : la base /lav/ se retrouve à toutes les formes et elle peut constituer un mot à elle seule : cf. l'énoncé impératif *Lav !* 'Lave !'. C'est une base libre. Le suffixe infinitif ou participial /-e/ est caractéristique de cette classe, qui correspond au 1<sup>er</sup> groupe de la grammaire traditionnelle ou encore à l'ensemble des verbes

<sup>136</sup> Proposition plutôt que phrase, car on sait qu'en FP *keskə* peut introduire une interrogation indirecte : cf. *ɜə=mə=dəmād keskə ty=fə* 'Je me demande qu'est-ce que tu fais' = FO et FP surveillé '... ce que tu fais'.

<sup>137</sup> La cooccurrence d'une particule et d'un mot interrogatifs n'est pas un phénomène rare : cf. japonais *Hon wa doko ni okimasita ka* {livre TOP où LOC avoir.mis Q} 'Où avez-vous mis le livre ?' (Kuwae 1983 : 158). L'haïtien, pourtant non sans attaches au FP, ne la permet pas, en revanche : cf. *Kisa (\*èske) ou bezwen ?* {quoi (\*Q) 2SG avoir.besoin} 'De quoi (est-ce que) tu as besoin ?' (Glaude 2012 : 195).

<sup>138</sup> Mes classes I et II correspondent aux 1<sup>er</sup>, 2<sup>e</sup> et 3<sup>e</sup> groupes traditionnels. Elles correspondent aussi aux macroclasses I et II de Kilani-Schoch & Dressler (2005 : 162 et suiv.), elles-mêmes subdivisées en une trentaine de microclasses. Le présent exposé n'exige pas un tel niveau de finesse, mais ne le récuse pas. Quiconque recherche un tableau exhaustif de la flexion verbale du FP devra compulsier cet ouvrage, de lecture, il est vrai, ardue pour des profanes. On lira aussi avec profit Blanche-Benveniste 2002.



dits « réguliers », parce que, connaissant la base unique et l'ensemble des suffixes, les formes en sont entièrement prédictibles. C'est, à peu de choses près, la seule classe ouverte (productive) susceptible d'accueillir de nouveaux membres.

Deux sortes de verbes de classe I paraissent dotées de plus d'une base. Il y a premièrement les verbes qui contiennent un schwa (*ə*), c.-à-d. une voyelle qui se manifeste comme /œ/ ou /E/ — c.-à-d. /ɛ/ ou /e/ selon les régions — ou pas du tout. Dans les verbes à base monosyllabique, cette voyelle unique se réalise /E/ en l'absence de désinence ou devant un suffixe à initiale consonantique : *ʒə=lev* 'je lève', *ʒə=levə-rɛ* 'je lèverai' ou 'je lèverais', *nu=levə-rjɔ̃* 'nous lèverions'. Le schwa qui apparaît à la fin de la base est phonotactique, sa fonction est de rendre la syllabation plus canonique, et sa réalisation est variable : rare — du moins dans ma variété « nordiste » de FP — dans *ʒə=levə-rɛ* /ʒlevɛ/, quasi constante dans *nu=levə-rjɔ̃* /nulɛvœkjɔ̃/. Quand la base est suivie d'un suffixe vocalique, le schwa est soumis à la fameuse « loi des trois consonnes » : *ʒə=ləv-ɛ* 'je levais' se prononce /ʒləvɛ/ ou /ʒœlvɛ/ — ou, bien évidemment, /ʒœlvœvɛ/ chez les francophones méridionaux qui prononcent davantage de schwas, ou bien dans toute occasion où parler clair paraît nécessaire pour une quelconque raison. *Nu=ləv-ɔ̃* n'a que deux réalisations : /nulvɔ̃/ et /nulœvɔ̃/.

Les verbes à base polysyllabique comme *aʃət-e* 'acheter', *epəl-e* 'épeler', *eləv-e* 'élever', *amən-e* 'amener', etc., où c'est la syllabe finale de la base qui comporte un schwa, diffèrent des monosyllabiques par quelques détails. Devant suffixe univocalique ce schwa est généralement muet en FP nordiste, mais souvent prononcé au sud : *ʒ=aʃət-ɛ* /ʒaʃ(œ)tɛ/ 'j'achetais', *ʒ=epəl-ɛ* /ʒep(œ)lɛ/ 'j'épelais', *ʒ=eləv-ɛ* /ʒel(œ)vɛ/ 'j'élèvais', *ʒ=amən-ɛ* /ʒam(œ)nɛ/ 'j'amenais'. Il lui correspond à nouveau un /E/ quand la base apparaît ou bien sans suffixe (*ʒ=aʃət* 'j'achète', *ʒ=epəl* 'j'épelle', *ʒ=eləv* 'j'élève', *ʒ=amən* 'j'amène') ou bien avec un suffixe à initiale consonantique : *ʒ=aʃətə-rɛ* 'j'achèterai', *ʒ=epələ-rɛ* 'j'épellerai', *ʒ=eləvə-rɛ* 'j'élèverai', *ʒ=amənə-rɛ* 'j'amènerai', et le schwa qui apparaît à la fin de la base est encore phonotactique et de réalisation variable : rare — pour moi — dans *ʒ=aʃətə-rɛ* où le groupe consonantique [tʁ] est sans problème (/ʒaʃɛtʁɛ/) ; plus fréquente dans *ʒ=epələ-rɛ* à cause du groupe /lʁ/ ; assez constante dans *vuz=aʃətə-rjɛ* 'vous achèteriez', *vuz=epələ-rjɛ* 'vous épelleriez', *vuz=eləvə-rjɛ* 'vous élèveriez', *vuz=amənə-rjɛ* 'vous amèneriez', réalisés /aʃɛtœrjɛ/, /epələœrjɛ/, /eləvœrjɛ/, /amənœrjɛ/.

Deux analyses sont envisageables pour les verbes de cette sorte. Ou bien on leur attribue deux bases, p.ex. /ləv-/ et /lEv-/, /aʃət-/ et /aʃEt-/, alternant comme esquissé ci-dessus. Ou bien on ne leur en attribue qu'une seule, pourvue d'une voyelle flottante (variablement réalisée) moyenne sous-spécifiée sur l'axe avant-arrière, qui, selon le contexte morphophonologique, se réalise antérieure, mi-haute ou mi-basse (ouverte ou fermée), non-arrondie ([ɛ] ou [e]), ou bien antérieure, mi-basse (ouverte), arrondie ([œ]), ou bien pas du tout. Comme souvent, le choix dépend d'hypothèses préalables quant à ce qui peut constituer une compétence naturelle pour les locuteurs natifs. La seconde analyse apparaît phonologiquement plausible et conforme à la diachronie des formes. La première est peut-être plus crédible d'un point de vue cognitif. Il ne m'appartient pas de trancher.<sup>139</sup>

L'autre apparente exception au caractère unibase de la classe I est celle des verbes dont la base se termine en /j/, p.ex. *nwaj-e* 'noyer', *pej-e* 'payer', *vej-e* 'veiller', *pjij-e* 'piller', *fuj-e* 'fouiller', *ãdœj-e* 'endeuiller', *ãnyij-e* 'ennuyer', etc. Final et précédé de la diphtongue /wa/ ou /qi/, ce /j/ disparaît systématiquement : cf. *ʒə=nwa* 'je noie' *ʒ=ãnyi* 'j'ennuie' vs. *vu=nwaj-e* 'vous noyez', *vuz=ãnyij-e* 'vous ennuyez', mais peut refaire surface à 3PL.PRS.INDIC : *i(l)=sə=nwa* 'ils se noient' et *i(l)=s=ãnyi* 'ils s'ennuient' prononcés /isnwaj/ et /isãnyij/ ; et à toutes les personnes du subjonctif : (*kə*) *ʒə=mə=nwa* '(que) je noie' et (*kə*) *ʒə=m=ãnyi* '(que) j'ennuie' prononcés /kœzmœnwaj/ et [kœzmãnyij] (cf. Morin

<sup>139</sup> Les verbes à plusieurs bases étaient plus nombreux en ancien français. Beaucoup ont été ramenés à l'unicité en moyen français.

2008).<sup>140</sup> Le /j/ des verbes du type de *pej-e* ‘payer’ est plus stable : *ʒə=pej* ‘je paye’ est commun (et reconnu par l’orthographe), tandis que *ʒə=pɛ* ‘je paie’ sonne plutôt obsolète, mais reste possible. A noter que l’instabilité ne touche pas de nombreux verbes où /j/ provient d’un ancien /l/ palatal (/ʎ/) comme *patruj-e* ‘patrouiller’ et *ʒə=patruj* ‘je patrouille’, *s=egaj-e* ‘s’égailer’ et *i(l)=s=egaj* ‘ils s’égailent’, etc. On comparera *s=egej-e* ‘s’égayer’ et *i(l)=s=egej* ‘ils s’égayent’ ~ *i(l)=s=ege* ‘ils s’égaiant’. Vu l’ignorance générale (et bien normale) en matière étymologique, la solution la plus réaliste est sans doute de considérer que tous ces verbes ont une seule base terminée par /j/, mais que celui-ci est instable chez certains. Ici, comme souvent, le FP se montre plus complexe, moins prévisible (plus entropique) que le FO qui, adossé à son orthographe historique, distingue mieux les séries.

Contrairement à ceux de la classe I (moyennant les possibles exceptions), les verbes de la classe II se fléchissent sur plusieurs bases. Celles-ci sont dites libres si elles peuvent constituer un mot-forme à elles seules, par exemple /part/ de *i(l)/ɛl=part* ‘ils/elles partent’ ; sinon elles sont liées, par exemple /byv-/ de *nu=byv-ɔ̃* ‘nous buvons’. Je l’indiquerai au moyen d’un tiret final. De même que pour les adjectifs, la différence est souvent entre une base courte, voire ultra-courte — par exemple /b-/, /bwa/, /par/, /dɔr/, /fini/, etc. — et une base longue comportant un appendice : /bwav/, /part/, /dɔrm/, /finis/, etc.

Trois suffixes d’infinitif caractérisent la classe II : *-r* (*perd-r* ‘perdre’, *ri-r* ‘rire’, *ekri-r* ‘écrire’, *ateri-r* ‘atterrir’), *-ir* (*part-ir* ‘partir’, *dɔrm-ir* ‘dormir’) et *-war* (*sav-war* ‘savoir’, *vul-war* ‘vouloir’). Quoique phonologiquement semblables à l’infinitif, les verbes du type de *ri-r*, *ekri-r*, *ateri-r* et ceux du type de *part-ir*, *dɔrm-ir* se distinguent en cela que /i/ appartient à la base dans les premiers, au suffixe dans les seconds : cf. *ʒə=ri* ‘je ris’, *ʒ=ekri* ‘j’écris’, *ʒ=ateri* ‘j’atterris’ vs. *ʒə=par* ‘je pars’, *ʒə=dɔr* ‘je dors’. Je ne souscrit pas à l’hypothèse selon laquelle, dans *part-ir* et *dorm-ir*, /i/ serait une voyelle « thématique », n’appartenant ni à la base ni au suffixe d’infinitif (cf. Kilani-Schoch & Dressler 2005 : 166). Selon moi, le fait décisif est que, hors l’infinitif, /i/ ne se retrouve dans aucune des bases sur lesquelles se fléchissent ces verbes, si bien qu’il paraît plus juste de le rattacher à la désinence d’infinitif, d’autant que, comme le reconnaissent les auteurs cités, l’existence de voyelles thématiques en français est douteuse. On vérifiera aisément que le même raisonnement justifie la différence de découpage entre *bwa-r* ‘boire’ et *vul-war* ‘vouloir’. Différence morphologique qui, comme on le voit, est corrélée à la présence ou non d’un *e* muet final en FO. Muet, cet *e* ne le fut pas toujours, et, aussi longtemps qu’il se prononça — guère plus tard que la fin du seizième siècle (cf. Gougenheim 1951 : 20) — le français eut une quatrième désinence d’infinitif, /-rə/, que le FO a conservée. Rien ne lui correspond en FP, à nouveau plus opaque.

La même opacité caractérise, au sein des verbes de classe II à base à finale /i/ et suffixe *-r* ou à suffixe *-ir*, la différence entre *ekri-r* et *part-ir*, d’une part, *ateri-r*, d’autre part. L’aspect de l’infinitif ne permet pas de prévoir que, dans les verbes du type de ce dernier, le nombre de bases alternantes se limite à deux et que l’alternance elle-même, /Xi-/ vs. /Xis-/, y est uniforme et prévisible, ce qu’elle n’est pas dans les premiers qui sont, par là, les seuls vrais « irréguliers ». Cela pose-t-il un problème pour les enfants qui acquièrent la langue ? Sans doute pas, s’il est vrai que l’infinitif n’est pas la forme la plus fréquente, en sorte qu’elles et ils entendent sans doute *nu=part-ɔ̃* ‘nous partons’ et *nuz=ateris-ɔ̃* ‘nous atterrissons’ plus tôt et davantage que *part-ir* et *ateri-r*.

Désignons ces deux sous-classes comme IIa pour *ekri-r* et *part-ir* et IIb pour *ateri-r*. Elles se distinguent par d’autres caractères que le nombre et la prédictibilité des bases. Premièrement, IIb reste marginalement productive alors que IIa est fermée : après *alyni-r* ‘alunir’ (1921) et *ameri-r* ‘amerrir’ (1928), *amarsi-r* ‘amarsir’ et son dérivé *amarsisaz* ‘amarsissage’ sont apparus récemment. Ira-t-on jusqu’à *avenyzi-r* ‘avénusir’, *atitani-r* ‘atitanir’, etc. ? Peu probable, mais qui sait ? Deuxièmement, les verbes dérivés d’adjectifs

<sup>140</sup> Ou /kʒmœnwaj/, /kʒmānujij/.

dénotant des propriétés physiques en relèvent le plus souvent : cf. *blãfi-r* ‘blanchir’, *mësi-r* ‘mincir’, *alurdi-r* ‘alourdir’, *vjoki-r* ‘vioquir’ (cf. Boyé 2000 : 23), etc., ce qui leur confère une prévisibilité spécifique, indépendamment de la question de savoir si la formation est productive ou non (cf. Kilani-Schoch & Dressler 2005 : 132-133) : si un verbe terminé en /-iɪ/ à l’infinitif se relie sémantiquement à un adjectif de cette sorte, il se fléchit conformément à IIb. Il y a donc peu de risques d’hésiter entre *nu=nwarsis-ǝ* ‘nous noirçissons’ et \**nu=nwars-ǝ* \*‘nous noirçons’, *ǝ=vjokis-ε* ‘je vioquissais’ et \**ǝ=vjok-ε* \*‘je vioquais’.<sup>141</sup>

Pour n’être pas globalement prévisibles, les alternances de IIa n’en sont pas pour autant dépourvues de régularités. Elles s’inscrivent en effet dans un « espace thématique » fait de relations relativement stables entre bases (Boyé 2000 ; Bonami & Boyé 2003).<sup>142</sup> Ainsi, il est remarquable que la base de la 3<sup>e</sup> personne du pluriel du présent de l’indicatif (6.PRS.INDIC) de *bwa-r*, /bwaɪ/, soit aussi celle des trois personnes du singulier et de la 3<sup>e</sup> personne du pluriel du présent du subjonctif (1-3, 6.PRS.SUBJ).<sup>143</sup> Cette correspondance est générale, et elle s’étend à tout le paradigme du présent du subjonctif quand la base de 4-5 ne diffère pas de celle de 6 au présent de l’indicatif : cf. *i(l)|εl=part* ‘ils|elles partent’ (*nu|vu=part-ǝ|e*) vs. *(kæ\_)ǝ=part* ‘(que) je parte’, *(kæ\_)nu=part-jǝ* ‘(que) nous partions’, *(k\_)i(l)|εl=part* ‘(qu’)il(s)|elle(s) parte(nt)’. Elle connaît des exceptions chaque fois que le subjonctif présente une base particulière : p.ex. *i(l)|εl=pæv* ‘ils|elles peuvent’, mais *(k\_)i(l)|εl=pwis* ‘(qu’)il(s)|elle(s) puisse(nt)’, *i(l)|εl=sav* ‘ils|elles savent’, mais *(k\_)i(l)|εl=saf* ‘(qu’)il(s)|elle(s) sache(nt)’, etc. En IIb, la base du subjonctif est toujours la base longue à toutes les personnes : *(kæ\_) ǝ=finis* ‘(que) je finisse’, *(kæ\_) nu=finis-jǝ* ‘(que) nous finissions’, *(k\_)i(l)|εl=finis* ‘(qu’)il(s)|elle(s) finisse(nt)’.

La notion d’espace thématique n’est pas sans relation avec cet outil grammatical plus familier et fort ancien — il remonte à l’Antiquité gréco-latine — que sont les temps primitifs, c.-à-d. les temps à partir desquels se forment tous les autres ou, pour le dire de façon un peu plus technique, les cases d’un paradigme dont le contenu permet de déduire le contenu de toutes les autres cases. Ainsi, pour BOIRE,<sup>144</sup> la case [TEMPS *présent*, MODE *indicatif*, PNG 1] réalisée *bwa* entraîne dans son sillage les cases [TEMPS *présent*, MODE *indicatif*, PN 2], [TEMPS *présent*, MODE *indicatif*, PNG 3], [TEMPS *futur*, PNG 1-6], [MODE *conditionnel*, PN 1-6] et [MODE *impératif*, PNG 2], réalisées sur la même base : cf. *ty=bwa* ‘tu bois’, *εl=bwa* ‘elle boit’, *ǝ=bwa-rε* ‘je boirai(s)’, *bwa !* ‘bois !’.<sup>145</sup> La 1<sup>ère</sup> personne du singulier du présent de l’indicatif peut donc être considérée comme un temps primitif qui rend prévisibles quinze cases du paradigme de BOIRE; avec un verbe de classe I comme *fãte* ‘chanter’ elle prédit toutes les cases contenant une forme synthétique, soit 47. Mais les alternances de bases dans les verbes autres que de classe I et IIb sont telles, en ampleur comme en imprévisibilité, que les présenter au moyen d’espaces thématiques plutôt que de temps primitifs, quoique les deux procédés soient largement équivalents, me paraît plus clair et, peut-être, plus conforme au savoir des sujets parlants, dont on peut supposer qu’ils ne « dérivent » pas les formes les unes

<sup>141</sup> Il existe au moins deux verbes mixtes. L’un est *fovir* ‘chavir’, IIb au singulier du présent : *i(l)=fovi* ‘il chavuit’, IIa partout ailleurs : 6 *i(l)=fov* ‘ils chavent’ (cf. Morin 1987a et Grevisse 1980 : §1637, mais pas Bescherelle, qui le range tout entier sous IIb, peut-être par confusion avec l’autre *fovir* ‘chavir’ qui signifie « devenir chauve »). Mais ce verbe, qui signifie « dresser les oreilles » en parlant d’un animal aux oreilles mobiles (âne, chouette, etc.), est si peu connu qu’on peut douter de son appartenance au FP, voire au FO commun. L’autre est *modir* ‘maudire’, IIb partout sauf au participe, IIa comme *di-r* : *modi* ‘maudit’, *modit* ‘maudite’. (Le participe des verbes IIb est épécène.) Le participe IIa de *benir* ‘bénir’ est en réalité un adjectif, guère employé que dans *o benit* ‘eau bénite’.

<sup>142</sup> « Thématiques » parce que « thèmes » s’emploie souvent comme un quasi synonyme de « base ».

<sup>143</sup> J’adopte la convention pratique de numéroter les personnes de 1 (1<sup>ère</sup> personne du singulier) à 6 (3<sup>e</sup> personne du pluriel).

<sup>144</sup> C’est-à-dire le lexème, noté en (petites) capitales et orthographe conventionnelle (voir chapitre 5).

<sup>145</sup> Ne sont indiqués que les traits appropriés à chaque case. Par exemple, le futur n’existe qu’à l’indicatif en français, il est donc inutile de noter le mode.

des autres, mais possèdent une certaine connaissance implicite, source de généralisations, des relations qui les unissent.

On notera l'existence en FP, exceptionnel en cela parmi les langues romanes, de bases consonantiques, manifestes au prétérit (passé simple), au passé du subjonctif et au participe de quelques verbes de la classe IIa : cf. *bwa-r* 'boire',  $\mathfrak{z}\partial=b-y$  'je bus', *nu=b-ym* 'nous bûmes', (*kə*) *nu=b-ysjð* '(que) nous bussions',  $\mathfrak{z}=\varepsilon_b-y$  'j'ai bu'.<sup>146</sup> Kilani-Schoch & Dressler (2005) récusent l'existence de telles bases. Leurs arguments pour ce faire sont toutefois plus de principe que factuels. Dans des verbes phonologiquement plus étoffés, -y est sans conteste un exposant ou une partie d'exposant du prétérit, du passé du subjonctif (dont la base est la même que celle du prétérit) et du participe : cf. *parət-r* 'paraître', *nu=par-ym* 'nous parûmes', (*kə*) *nu=par-ysjð* '(que) nous parussions',  $\mathfrak{z}=\varepsilon_{par}-y$  'j'ai paru'. L'alternative serait de supposer une base /by/ exceptionnellement porteuse par défaut de ces trois valeurs de TAM qui, à la différence du présent, peuvent difficilement être considérées comme sémantiquement non-marquées. L'hypothèse me paraît plus coûteuse, moins plausible que celle d'une base consonantique, dont on ne voit pas qu'elle ait rien qui choque l'entendement. Comme *bwa-r*, il y a *krwa-r* 'croire' ( $\mathfrak{z}\partial=kr-y$  'je crus'), *pløv-war* (*il=a\_pl-y* 'il a plu'), *fwa-r* 'choir' ( $\mathfrak{z}\partial=f-y$  'je chus'), *vwa-r* 'voir' ( $\mathfrak{z}\partial=v-i$  'je vis'). Mais les verbes incluant une diphtongue /wa/ ne sont pas seuls dans ce cas : cf. *plē-r* 'plaire' et *sa m=a\_pl-y* 'ça m'a plu', *ri-re* 'rire' et *nu=r-im* 'nous rîmes', etc.

Ces bases consonantiques sont le résultat de processus généraux de chute des consonnes intervocaliques et de réduction des hiatus qui ont très tôt affecté l'ancien français, altérant les formes héritées du latin tardif à un degré bien supérieur à ce qu'ont connu les autres langues romanes. Du point de vue diachronique, il est établi qu'à l'origine se trouve l'instabilité de l'accent tonique dans les paradigmes du parfait latin, par exemple, en latin tardif du nord de la Gaule tel que reconstitué par Moignet (1988 : 74), pour *videre* 'voir', *vidi* 'je vis' et *vedisti* 'tu vis', celui-là forme forte accentuée sur la base, celui-ci forme faible accentuée sur la désinence. Cela donne en ancien français, pour *veoir* /və'wɛr/ 'voir', *vi* 'je vis' et *veis* /və'is/ 'tu vis', où la personne 1 se réduit à l'évidence à la base étymologique (Moignet 1988 : 73). On peut, je l'admets malgré mes réticences à mêler diachronie et synchronie, en tirer argument pour réfuter l'hypothèse de bases consonantiques. Le prix, je l'ai dit, une anomalie sémantique, surtout sensible au participe, des trois la seule forme vraiment vivante en FP. Mais peut-être est-elle insignifiante quant au plus important, l'acquisition de la langue.

Pourquoi *bwa-r*, *vwa-r*, etc. mais *pløv-war*, *vul-war* 'vouloir', etc. ? D'un point de vue synchronique, le motif de la différence de découpage est clair : dans le second groupe, la diphtongue /wa/ n'apparaît nulle part ailleurs qu'à l'infinitif. Il est donc judicieux de la rattacher à la désinence de celui-ci plutôt qu'à la base.

Les grammaires qui ne prennent en compte que l'écrit — c'est-à-dire presque toutes — passent en outre à côté d'un détail formel qui contribue peut-être à la reconnaissance de la répartition des verbes entre les trois classes. C'est, me semble-t-il, un fait avéré qu'en FP les verbes dont les trois personnes du singulier et la 3<sup>e</sup> personne du pluriel du présent de l'indicatif — identiques (syncrétiques) à l'exception de quelques verbes isolés — se terminent par une consonne quelconque, sauf /t/, appartiennent à la classe I : cf.  $\mathfrak{z}\partial=t\tilde{\delta}b$  'je tombe',  $\mathfrak{z}\partial=sud$  'je soude',  $\mathfrak{z}\partial=kwaf$  'je coiffe',  $\mathfrak{z}\partial=læg$  'je lègue',  $\mathfrak{z}\partial=m\tilde{a}z$  'je mange',  $\mathfrak{z}\partial=fik$  'je chique',  $\mathfrak{z}=ap\tilde{e}l$  'j'appelle',  $\mathfrak{z}\partial=m=\tilde{a}rym$  'je m'enrhume',  $\mathfrak{z}\partial=d\tilde{o}n$  'je donne',  $\mathfrak{z}=atr\tilde{a}p$  'j'attrape',  $\mathfrak{z}\partial=les$  'je laisse',  $\mathfrak{z}\partial=f\tilde{a}t$  'je chante',  $\mathfrak{z}\partial=r\tilde{e}v$  'je rêve',  $\mathfrak{z}=\tilde{o}z$  'j'ose'. Coextensives

---

<sup>146</sup> Le prétérit n'est certes plus guère en usage en FP. Mais sa connaissance passive est entretenue, non sans lacunes, par son emploi constant dans la fiction littéraire, y compris celle destinée aux enfants. On ne saurait donc l'exclure du FP. Le sort du subjonctif passé semble scellé, en revanche, sauf comme réserve de plaisanteries (voir chapitre 10).

de la base, ces formes expriment donc leur valeur de TAM par défaut. Les verbes à finale /r/ aux mêmes formes relèvent ou bien de la classe I ( $\mathcal{Z}$ =*ader* ‘j’adhère’), ou bien de la classe IIa ( $\mathcal{Z}$ =*par* ‘je pars’).<sup>147</sup> Les verbes à finale vocalique entrent dans les trois classes : cf.  $\mathcal{Z}$ =*brwa* ‘je broie’ (I),  $\mathcal{Z}$ =*krwa* ‘je crois’ (IIa),  $\mathcal{Z}$ =*prā* ‘je prends’ (IIa),  $\mathcal{Z}$ =*kre* ‘je crée’ (I),  $\mathcal{Z}$ =*tre* ‘je trais’ (IIa),  $\mathcal{Z}$ =*krē* ‘je crains’ (IIa),  $\mathcal{Z}$ =*ofisi* ‘j’officie’ (I),  $\mathcal{Z}$ =*ateri* ‘j’atterris’ (IIb),  $\mathcal{Z}$ =*kōfi* ‘je confie’ (I) ou ‘je confis’ (IIa),  $\mathcal{Z}$ =*repō* ‘je réponds’ (IIa),  $\mathcal{Z}$ =*mō* ‘je meus’ (IIa),  $\mathcal{Z}$ =*kōspy* ‘je conspue’ (I),  $\mathcal{Z}$ =*lu* ‘je loue’ (I),  $\mathcal{Z}$ =*ku* ‘je couds’ (IIa).

L’histoire explique ces corrélations. Après les régularisations survenues pendant la dernière période de l’ancien français (quatorzième-quinzième siècles), les PNG 1, 2 et 3 du présent de l’indicatif des verbes de classe I (verbes en *-er/-ier*) se sont trouvées porteuses d’un schwa final (suivi de /s/ à PNG 2) qui ne caractérisait auparavant que les PNG 2 et 3 et les verbes à base terminée par un groupe /Cr/ : on conjugait *chant* /ʃã<sup>n</sup>t/, *chantes* /ʃã<sup>n</sup>təs/, *chante* /ʃã<sup>n</sup>tə/ et *entre* /ã<sup>n</sup>trə/, *entres* /ã<sup>n</sup>trəs/, *entre* /ã<sup>n</sup>trə/ (cf. Moignet 1988 ; Fouché 1967). Comme on sait, ce schwa s’est prononcé plus ou moins nettement jusqu’au dix-septième siècle avant de s’amuir. De ce fait, les consonnes d’attaque des syllabes finales de ces formes — par exemple le /t/ de  $\mathcal{Z}$ =*ʃã<sup>n</sup>tə* — ont échappé à l’effacement général des consonnes finales en moyen français (seizième siècle) (cf. Gougenheim 1951), avant de devenir elles-mêmes finales une fois /ə/ disparu de la prononciation — mais pas de l’orthographe. Rien, en revanche, ne permettait une généralisation de schwa dans les verbes IIa dont une unique consonne terminait la base : *sentir* a continué de se conjuguer *sent* /sã<sup>n</sup>t/ ‘je sens’, *senz* /sã<sup>n</sup>ts~sã<sup>n</sup>s/ ‘tu sens’, *sent* /sã<sup>n</sup>t/ ‘elle/il sent’. Toutes ces consonnes restées finales sont tombées en moyen français, avec pour résultat des formes à finale vocalique, sauf lorsque la finale était un /r/ précédé d’une consonne et de ce fait protégé par un schwa comme dans *ouvrir*, *uevre* /œvrə/ ‘j’ouvre’, *uevres* /œvrəs/ ‘tu ouvres’, *uevre* /œvrə/ ‘elle/il ouvre’. (La survie de /r/ dans  $\mathcal{Z}$ =*kur* ‘je cours’,  $\mathcal{Z}$ =*mør* ‘je mords’, etc. est sans doute analogique.) La finale de la base courte des verbes IIb a, quant à elle, toujours été /i/.

On aboutit ainsi en FP à une division phonologique assez nette : d’une part des formes à finale consonantique toujours intégrées à des paradigmes de verbes réguliers, de flexion entièrement prédictible, majoritaires et multipliables ; d’autre part, sans doute moins nombreuses, des formes à finale vocalique, laquelle ne dit rien de leur rattachement à l’une ou l’autre des trois classes. Il est vrai que cette division ne se montre qu’à condition de procéder à l’inverse de la démarche grammaticale habituelle, des formes perçues au type flexionnel plutôt que du type flexionnel aux formes. Du point de vue de l’acquisition, ce parcours paraît plus naturel. Lui correspond-t-il quelque chose dans la réalité ? Une étude expérimentale le dira peut-être.

A quelque classe qu’ils appartiennent, les verbes présentent des formes finies référant à des évenances (événements et états) délimitables (en principe sinon de fait) et des formes non-finies qui n’ont pas cette propriété. Les premières se fléchissent pour les propriétés de temps-aspect-mode (TAM), personne-nombre-genre (PNG), voix (Vx) active ou passive et polarité (POL) positive ou négative. Les secondes (infinitif, participe, gérondif) ne connaissent que le temps, la voix, la polarité et, pour le participe, le genre. TAM et PNG sont souvent cumulés au sein d’une forme unique indécomposable. Ainsi, dans *ʃãt-ε*, l’exposant *-ε* exprime simultanément la valeur de TAM imparfait (ou passé imperfectif) de l’indicatif et les valeurs de PNG 1-3, 6 :  $\mathcal{Z}$ |*ty|εl*=*ʃãt-ε* ‘je|tu chantais, elle chantait, elles chantaient’. La flexion procède selon les quatre modes dégagés au chapitre 5 : synthétique et suffixal pour TAM et, avec certaines formes seulement, PNG (*ʃãt-ε*, *ʃãt-jō* ‘chantions’, *ekrit* ‘écrite’) ; semi-analytique et clitique pour PNG et la polarité négative (*εl(=ne=)ʃãt-ε=pa* ‘elle ne chantait pas’) ; analytique (périphrastique) pour TAM et la voix passive ((*sə=ʃã*) *a\_ete\_ʃãte* ‘(ce

<sup>147</sup> Ce /r/ peut faire partie d’un groupe /Cr/ : cf.  $\mathcal{Z}$ =*ãtr* ‘j’entre’ (I),  $\mathcal{Z}$ =*uvr* ‘j’ouvre’ (IIa). Les verbes à /j/ final relèvent aussi de ces deux classes : cf.  $\mathcal{Z}$ =*m=evēj* ‘je m’éveille’ (I),  $\mathcal{Z}$ =*kæj* ‘je cueille’ (IIa).

chant) a été chanté’) ; défaut (forme réduite à la base) pour TAM-PNG ( $ʒə|ty|ɛl=f\tilde{a}t$ ). Tout cela sera examiné en détail dans les chapitres suivants.

Un dernier mot : le système — si on peut le nommer ainsi — d’alternance des bases, pour ne concerner qu’une minorité de verbes, mais dont beaucoup figurent parmi les plus fréquemment employés, est évidemment d’acquisition difficile, vu son degré élevé d’imprévisibilité. Il donne donc lieu à des « fautes » telles que \**pēd-ε* au lieu de *pep-ε* ‘peignais’ à l’imparfait de *pēd-r* ‘peindre’ (Blanche-Benveniste 1997 : 145). L’analogie, la régularisation y tiennent un rôle majeur. Mon but dans cet ouvrage n’étant pas de les relever, encore moins de les déplorer, je renvoie aux ouvrages qui en font l’étude, dont celui de l’auteur cité à l’instant, mais aussi Frei (1929/2007) et Kilani-Schoch & Dressler (2005 : 187 et suiv.).

Je donne à présent, à titre d’exemple, les tableaux de répartition des bases pour un verbe de classe IIa, *vul-war* ‘vouloir’, et un verbe de classe IIb, *ateri-r* ‘atterrir’. (En faire autant pour un verbe régulier de classe I serait évidemment sans intérêt.) La flexion du premier requiert cinq bases : /*vø*/, /*vul-*/, /*vœl*/, /*vœj*/, /*vud-*/, celle du second, deux, /*ateri*/ et /*ateris*/ :<sup>148</sup>

	/vø/	/vul-/	/vœl/	/vœj/	/vud-/
présent	1-3	4-5	6		
futur					1-6
conditionnel					1-6
imparfait		1-6			
prétérit		1-6			
subjonctif présent		4-5		1-3, 6	
subjonctif passé		1-6			
impératif				2, 5	
participe		✓			
gérondif		✓			
infinitif		✓			

Cf.  $ʒə|ty|i(l)|ɛl=vø$  ‘je/tu/il/elle veut/veut’,  $nu=vul\tilde{ɔ}$  ‘nous voulons’,  $i(l)|ɛl=vœl$  ‘ils/elles veulent’,  $ʒə=vud-rɛ$  ‘je voudrai’, etc.,  $ʒə=vud-rɛ$  ‘je voudrais’, etc.,  $vœj !$  ‘veuille’, *vul-war* ‘vouloir’, ( $kæ$ )  $ʒə=vœj$  ‘(que) je veuille’,  $ʒə=vul-y$  ‘je voulus’, etc., ( $kæ$ )  $ʒə=vul-ys$  ‘(que) je voulusse’, etc., *voul-y* ‘voulu.e(s)’, ( $kæ$ )  $nu=vul-j\tilde{ɔ}$  ‘(que) nous voulions’,  $ʒə=vul-ε$  ‘je voulais’, *vul-ã* ‘voulant’.

	/ateri/	/ateris/
présent	1-3	4-6
futur	1-6	
conditionnel	1-6	
imparfait		1-6
prétérit	1-6	
subjonctif présent		1-6
subjonctif passé	1-6	
impératif	2	5
participe	✓	
gérondif		✓

<sup>148</sup> J’appelle gérondif le participe présent de la grammaire scolaire.

infinitif	✓	
-----------	---	--

Cf. *L=avjḑ ateri* ‘L’avion atterrit’, *L=avjḑ ateri-ra* ‘L’avion atterrira’, *L=avjḑ ateri-re* ‘L’avion atterrira’, *Nuz=ateri-m* ‘Nous atterrîmes’, (*il=y\_faly kœ*) *nuz=ateri-sjḑ* ‘(il eût fallu que) nous atterrissons’, *Ateri !* ‘Atterris !’, *L=avjḑ a\_ateri* ‘L’avion a atterri’, *Nuz=ateris-ḑ* ‘Nous atterrissons’, *Nuz=ateris-jḑ* ‘Nous atterrissons’, (*i(l)=fo kœ*) *nuz=ateris-jḑ* ‘(il faut que) nous atterrissons’, *Ateris-e !* ‘Atterrissez !’, *ateris-ã* ‘atterrissant’, *ateri-r* ‘atterrir’.

La comparaison des deux tableaux montre que seul le présent de l’indicatif (et du subjonctif présent pour *vul-war*) requiert plus d’une base, tandis que les autres formes utilisent l’une ou l’autre base dans tout leur paradigme. On verra qu’il y a là une régularité assez générale.

## 10. Les formes finies du verbe

Le terme « forme finie » n’est pas traditionnel. On risque d’y voir un anglicisme, transposition de *finite form*. En fait, il s’agit plutôt de la traduction du latin *verbum finitum* que Marouzeau (1961 : 95) définit ainsi : « Ensemble des formes verbales personnelles susceptibles d’exprimer le rapport de l’action à un sujet déterminé ». Autrement dit, *fāt-ḑ* ‘chantons’ est fini (*finitum*) parce que la forme même attribue le chant à NOUS, un ensemble de personnes incluant la locutrice ou le locuteur. Les formes non-finies, infinitif, participe et gérondif, en revanche, n’ont pas cette propriété. Il est vrai qu’on peut dire *mwa, fāt-e sa, zame !* ‘Moi, chanter ça, jamais !’. Sauf que *\*zə=fāt-e sa* ‘je chanter ça’ est agrammatical et que remplacer *mwa* ‘moi’ par *nu* ‘nous’ n’a aucun effet sur la forme du verbe — et l’infinitif « remplacer » de cette phrase n’attribue l’action à personne en particulier. Sont donc finies en FP toutes les formes à l’exception des trois citées plus haut.

Comme expliqué à la fin du chapitre précédent, les procédés flexionnels générant des formes finies (et aussi non-finies) sont : (a) la suffixation (flexion synthétique), (b) la cliticisation (flexion semi-analytique), (c) la périphrase (flexion analytique), (d) le défaut (thème nu, éventuellement supplétif). Ces quatre procédés se ramènent à trois si l’on considère le défaut (la flexion par absence de marque) comme le cas limite de la synthèse, tel que n’affixer rien continue à faire sens. Outre la personne-nombre et, à la 3<sup>e</sup> personne, le genre du sujet, les formes finies expriment des valeurs combinables (et le plus souvent combinées) de temps, d’aspect et de mode (TAM). Les temps sont le présent (réel ou atemporel), le passé, le futur ; les aspects, le perfectif, l’imperfectif ; les modes, l’indicatif, le subjonctif, l’impératif. Je garde le terme traditionnel « imparfait » pour le passé imperfectif (*zə=fāt-e* ‘je chantais’). Mais je préfère « prétérit » à « passé simple » pour désigner le passé perfectif (*nu=fāt-am* ‘nous chantâmes’, *nuz=avḑ\_fāt-e* ‘nous avons chanté’). Le conditionnel fait problème : temps ou mode ? Son emploi en contexte hypothétique ainsi que son fréquent emploi médiatif — information rapportée sans engagement de qui la rapporte, p.ex. *La=zokḑd sə-re(t)\_ēn=əm* ‘La Joconde serait un homme’ —, l’un et l’autre sans doute historiquement corrélés, incitent à le considérer comme un mode. Temporelle, en revanche — et conforme à son étymologie — est sa valeur de futur dans un contexte de passé : *Lə=zardinje esper-e k\_i(l)=plæv-re* ‘Le jardinier espérait qu’il pleuvrait’.<sup>149</sup> En synchronie, il ne semble pas qu’on puisse faire mieux qu’enregistrer cette dualité.

La flexion TAM se réalise par les procédés (a/d) et (c) : cf. (*zə=*)*fāt* ‘(je) chante’ (présent de l’indicatif, synthétique-défaut) ; (*vu=*)*fāt-e* ‘(vous) chantez’ (présent de l’indicatif, synthétique suffixal) ; (*z=*)*ε\_fāt-e* ‘(j’)ai chanté’ (prétérit périphrastique). La flexion PNG se

<sup>149</sup> De même que le futur procède d’une périphrase infinitif + présent du latin *habere* ‘avoir’, le conditionnel a pour origine une périphrase infinitif + imparfait de *habere*.

réalise par (a/d) et (b) : cf.  $zə=fāt$  (présent de l'indicatif 1, semi-analytique + synthétique-défaut) ;  $vu=fāt-e$  (présent de l'indicatif 5, semi-analytique + synthétique suffixal). Après bien d'autres, j'analyse donc les pronoms faibles  $zə$  'je',  $vu$  'vous', etc. comme des composantes des formes verbales finies indexant les traits de personne et de nombre du sujet desdites formes, équivalents clitiques d'affixes tels que le  $-e$  de  $vu=fāt-e$  'vous chantez', avec lesquels ils cooccurrent.

La 3<sup>e</sup> personne (3 et 6) appelle un développement particulier, car elle oblige à expliciter la notion d'accord, jusqu'à présent implicite. Par accord, on le sait, on entend le partage visible (par un quelconque procédé morphologique) de certaines propriétés par les éléments de certaines constructions (voir chapitre 5 et Creissels 2006a : 24 et suiv.). Or, aux personnes 3 et 6, le clitique personnel peut manquer en présence d'un SN sujet : cf.  $me=vwazē m=ābet$  'Mes voisins m'embêtent' où l'accord PNG 6 n'est marqué que par défaut, la base nue des verbes de classe I étant compatible avec cette valeur, ainsi qu'avec 1-3 comme on sait (et voir plus loin). (On comparera avec  $me=vwazē sō(t) ābetā$  'Mes voisins sont embêtants', où l'accord PNG 6 est explicite.) Mais l'on dit aussi  $me=vwazē i(l)=m=ābet$  'Mes voisins, ils m'embêtent' sans que le SN sujet en reçoive forcément une valeur pragmatique particulière, emphase ou thématization (indication d'avoir déjà été mentionné dans le discours en cours), si bien que les deux expressions sont souvent interchangeables.<sup>150</sup> Autrement dit, l'accord sujet-verbe aux personnes 3 et 6 en FP se réalise ou bien seulement par le procédé synthétique-défaut, ou bien par celui-ci *et* le procédé semi-analytique. C'est le contour prosodique de l'énoncé — dont la courbe devrait toujours être intégrée à la description grammaticale — qui indique comment interpréter le SN initial : sujet ou élément détaché. Et l'on parlera d'accord dans le premier cas, mais plutôt de concordance dans le second.

En l'absence de SN sujet, en revanche, le clitique personnel est évidemment obligatoire et  $*m=ābet$  \*'M'embête(nt)' n'est pas, à la différence de l'occitan  $m'embestia(n)$ , une phrase bien formée — sauf énoncé elliptique, déviant et senti comme tel, mais acceptable sous certaines conditions. On en conclut souvent que le FP (et le FO) diffère de la plupart des autres langues romanes du fait de n'être pas, comme elles, une langue « à sujet nul », c.-à-d. telle que l'expression des traits PNG du sujet peut être prise entièrement en charge par la flexion synthétique.<sup>151</sup> On attribue souvent cette divergence à la relative pauvreté de ladite flexion en FP : les verbes de classe I, de loin les plus nombreux, ne présentent au présent de l'indicatif que trois formes distinctes : 1-3, 6 ( $fāt$ ), 4 ( $fāt-ō$ ) et 5 ( $fāt-e$ ), voire deux quand  $ō=fāt$  est employé au sens de  $nu=fāt-ō$  (voir ci-dessous) — tandis qu'un équivalent occitan, espagnol, etc. en présente six.

Le raisonnement ne tient guère devant la réalité. Le paradigme du présent de l'indicatif d'un verbe de classe I en ancien français (douzième-treizième siècles) distingue formellement cinq formes sur six (les mêmes qu'en FO actuel) :  $chante$  /ʃãntə/ (1,3),  $chantes$  /ʃãntəs/ (2),  $chantons$  /ʃãntōs/ (4),  $chantez$  /ʃãntets/ (5),  $chantent$  /ʃãntən(t)/ (6). Pourtant,

« Si, en très ancien français, l'emploi du pronom reste rare et généralement expressif, répondant à une intention d'insistance ou d'opposition, il devient par la suite plus fréquent, puis courant et normal, sans qu'aucune nuance d'insistance stylistique se laisse percevoir. Il devient la marque de la personne sujet. Dans la prose du début du XIII<sup>e</sup> siècle, c'est l'absence du pronom sujet qui devient relativement rare. » (Moignet 1988 : 128)

<sup>150</sup> Comme on voit, la traduction en FO requiert une virgule entre *Mes voisins* et *ils m'embêtent*, indice d'une construction obligatoirement « détachée », pragmatiquement chargée, à la différence de sa contrepartie FP. Il se peut qu'il en aille de même dans les variétés les plus standard du FP, s'il est vrai qu'on y observe toujours une pause plus ou moins perceptible là où l'écrit a une virgule. Cela reste à vérifier.

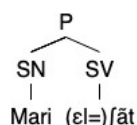
<sup>151</sup> En fait PN, car la flexion synthétique ne mentionne pas le genre dans les langues indo-européennes, à la différence des langues sémitiques, par exemple. Font aussi exception les autres langues d'oïl et les langues du nord de l'Italie, lombard, vénitien, etc.



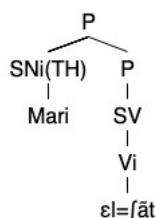
Or, les pronoms sujets de l'ancien français (*je~jo~g(i)é, tu, il/ele, no(u)s, vo(u)s, il/eles*) ne sont nullement des clitiques, ce sont des mots pleins que d'autres mots pleins peuvent séparer du verbe comme dans la tournure figée « Je soussigné déclare... ». En fait — et cela ne saurait nous surprendre, vu la force de l'influence germanique — l'ancien français ressemble sur ce point à l'allemand où le paradigme du présent de l'indicatif distingue formellement quatre personnes sur six, soit pour *singen* 'chanter', *singe* (1), *singst* (2), *singt* (3,5), *singen* (4,6) ; les pronoms sujets (*ich, du, er/sie/es, wir, ihr, sie*) sont néanmoins obligatoires, sauf à la 3<sup>e</sup> personne en présence d'un SN sujet ; lesdits pronoms sont des mots pleins.

L'évolution des pronoms sujets autonomes vers des clitiques personnels, encore inachevée aux seizième et dix-septième siècles (cf. Gougenheim 1951 : 68 ; Spillebout 1985 : 143), est ce qui a peu à peu distingué le FP, ancien et moderne, de l'allemand (et du FO). C'est là que la réduction de la flexion synthétique due à la chute des consonnes finales qui touche l'ancien français tardif et le moyen français a probablement joué un rôle, pour autant que cette réduction s'est vue compensée par la montée du marquage semi-analytique des propriétés PNG, lequel a sans doute aussi été favorisé par la débilite phonétique des pronoms.<sup>152</sup> Mais il en a résulté une conséquence à première vue paradoxale, à savoir que c'est alors que le FP est (re)devenu une langue « à sujet nul », où *i(l)=fāt* 'il chante' constitue un énoncé complet bien formé à l'instar de l'occitan *canta*. Les procédés de marquage différent, mais le résultat est le même.

On notera toutefois que le FP n'est pas allé aussi loin que les langues nord-italiennes où la flexion PNG semi-analytique semble à peu près généralisée : cf. vénitien *La Maria la canta* 'Marie chante', *la* clitique personnel 3 féminin. Le FP offre le choix entre *Mari fāt* et *Mari el=fāt*, dont l'arborescence ci-dessous illustre la communauté de structure :



Différente, en revanche, sera la structure de *Mari, el=fāt* lorsque la prosodie fait entendre que le SN *Mari*, thématif, mis en relief, n'est pas le sujet syntaxique de *el=fāt* :<sup>153</sup>



Le verbe étant alors privé de SN sujet dans son domaine phrastique (P inférieur), le clitique personnel sujet est obligatoire. C'est lui seul qui permet la coindexation (identité de référence) quant aux traits PNG entre le verbe fléchi ( $V_i$ ) et le SN thématif ( $SN_i$ ).

Des enquêtes sociolinguistiques de grande envergure seraient nécessaires pour déterminer s'il est vrai que le FP évolue vers une réalisation constante et généralisée des clitiques personnels sujets, comme on le dit souvent, mais sans vrai fondement empirique, du français « populaire » ; ou bien si l'alternance perdure, auquel cas il conviendrait de découvrir s'il s'agit d'une variation conditionnée (et par quels facteurs) ou bien entièrement libre.<sup>154</sup> En l'état actuel, le premier embranchement de la seconde branche de l'alternance (alternance conditionnée) me paraît le plus probable. C'est du reste pourquoi je continue à parler de *clitiques*, définition qui préserve un certain degré d'autonomie lexicale, et évite le terme d'indices — lequel est en revanche approprié pour une langue comme le swahili, langue

<sup>152</sup> La réalisation [i] de *il(s)* est attestée dès le moyen français et probablement plus ancienne (cf. Zink 1990 : 15).

<sup>153</sup> P = phrase, TH = thème (aussi appelé « topique »).

<sup>154</sup> Fait évidemment exception l'impératif, dont la valeur PNG (2) n'est jamais exprimée par un clitique.

bantoue parlée nativement ou comme langue seconde par quelques 100 millions de personnes en Afrique de l'Est. Là, les marques PNG sont de fait des préfixes obligatoires en toutes circonstances : cf. *Nimefika* {ni<sub>1,S</sub>-me<sub>PARFAIT</sub>-fika} 'Je suis arrivé.e', *Mgeni amefika* {mgeni a<sub>3,S</sub>-me<sub>PARFAIT</sub>-fika} 'Le visiteur / La visiteuse est arrivé.e' (Perrott 1951). Autrement dit, la flexion du swahili n'est pas (ou plus) semi-analytique, mais synthétique préfixale. On ne peut exclure que le FP en vienne un jour à adopter ce système, voire l'ait déjà adopté dans telles variétés, mais rien n'indique qu'il en aille présentement ainsi pour la majorité de ses locutrices et locuteurs. L'étude trop négligée ou trop confidentielle des autres langues d'oïl fournirait à ce propos d'intéressantes données comparatives. Ainsi, à propos du parler champenois de Gespunsart, Descusses (1986 : 96) écrit : « ... les troisièmes personnes... sont souvent, en patois comme en français populaire, en situation de redondance avec le substantif qu'ils sont censés reprendre », ce dont elle donne, entre autres, l'exemple de *la butik l=ε tRāsfo:Rmej La=butik el=ε trāsforme* 'La boutique (elle) est transformée'.<sup>155</sup>

Par défaut proclitiques conformément à la typologie SVO (sujet < verbe < objet) du FP, les clitiques personnels sujets sont enclitiques (attachés à la fin de la forme verbale) dans certaines constructions dites « inversées », en particulier les phrases interrogatives : *M=em=ty ?* 'M'aimes-tu ?'.<sup>156</sup> Dans ce cas et aux personnes 3 et 6 dont les indices *i(l)* et *el* ont une initiale vocalique, un [t] de liaison apparaît. Dans quelques formes, ce [t] se rattache historiquement à la terminaison : *M=em•t=el ?* 'M'aiment-elles ?' ; mais bien plus souvent non : *M=em•t=el ?* 'M'aime-t-elle ?' — et l'orthographe en tient compte. C'est pourquoi l'analyse en consonne de liaison (paragogique) apparaît comme la seule recevable en synchronie.

En toute rigueur, analyser les pronoms faibles comme des clitiques personnels devrait conduire à inclure les éventuels pronoms faibles objets dans les formes fléchies, si bien que, p.ex., *zə=lə=luj=ε dən-e* 'je le lui ai donné' remplit une case du paradigme complet de *don-e* 'donner'. Davantage encore que les sujets, les clitiques personnels objets peuvent s'absenter en présence de SN objets. Ils sont toutefois fréquents du fait de la propension du FP à produire des énoncés tels que *mari<sub>3</sub> lə=bukē<sub>2</sub> sɔ̃=kopē<sub>1</sub> i(l)<sub>1</sub>=lə<sub>2</sub>=luj<sub>3</sub>=a rādy* 'Marie, le bouquin, son copain, il le lui a rendu' (FO 'Son copain a rendu le bouquin à Marie'), typologiquement proches de ce qu'on observe dans les langues dites « polysynthétiques » telles les langues athabaskanées en Amérique du Nord, les langues bantoues en Afrique, etc. Les SN argumentaux extériorisés font fonction de thèmes repris par les clitiques personnels intégrés au SV comme indiqué par les chiffres souscrits. Ce qui autorise le rapprochement avec les langues polysynthétiques — qu'il ne convient sans doute pas d'exagérer — c'est le fait que cette thématization (ou topicalisation) est banale et ne requiert ni contexte discursif particulier, ni intonation spéciale : on peut énoncer la version FP de l'exemple dans les mêmes circonstances et avec la même prosodie qu'on en énoncerait la version plus proche de la syntaxe du FO.<sup>157</sup>

### 10.1. Les formes finies (partiellement) synthétiques

<sup>155</sup> La transcription en italiques gras est celle de l'autrice.

<sup>156</sup> Quoique plus rare que *eskə ty=m=em ?* 'Est-ce que tu m'aimes ?' et *ty=m=em ?* 'Tu m'aimes ?', ce tour appartient sans conteste au FP. Il est ordinaire au Québec.

<sup>157</sup> En swahili, l'indice personnel objet n'est régulièrement associé à un SN objet que si celui-ci est défini : cf. *Umekileta kitabu ?* {u<sub>2,S</sub>-me<sub>PARFAIT</sub>-ki<sub>3,O</sub>-leta kitabu} 'As-tu apporté le livre ?', vs. *Umeleta kitabu ?* 'As-tu apporté un livre ?' (Perrott 1951). A comparer avec FP *ty=l=a aporte lə/\*ε=bukē ?* 'Tu l'as apporté, le/\*un bouquin'. Clitiques ou indices, les exposants pronominaux sont intrinsèquement définis, cela explique la correspondance. A noter que l'ordre des arguments thématiques est relativement libre — cf. *mari sɔ̃=kopē lə=bukē...* — si ce n'est que l'objet indirect doit être marqué s'il n'est pas initial : *sɔ̃=kopē lə=bukē a=mari ...* L'ordre des clitiques est fixe, en revanche. On voit là la différence entre syntaxe et morphologie.

Comme on vient de le voir, la synthèse des formes finies se combine avec la semi-analyticité de la flexion PNG, d'où l'adverbe entre parenthèses du titre de cette section. Seule la dimension synthétique des formes nous occupera à présent.

Sans doute est-ce un parti raisonnable que de mettre à part trois verbes qu'on dira « anomaux », car ils n'appartiennent ni à la classe I ni aux classes IIa et IIb : *et-r* 'être', *av-war* 'avoir' et *al-e* 'aller'.<sup>158</sup> Le premier ne comporte pas moins de neuf bases supplétives, le deuxième sept, le troisième six :<sup>159</sup>

	/εt/	/sqi/	/ε/	/sɔm/	/sɔ̃/	/swa/	/swaj/	/sə-/	/f-/
présent	5	1	2-3	4	6				
futur								1-6	
conditionnel								1-6	
imparfait	1-6								
prétérit									1-6
subjonctif présent						1-6	(1-2, 6)		
subjonctif passé									1-6
impératif						2	5		
participe	✓								
gérondif	✓								
infinitif	✓								

	/av-/	/ε/	/a/	/ɔ̃/	/o-/	/y/	/εj/
présent	4-5	1	2-3	6			
futur					1-6		
conditionnel					1-6		
imparfait	1-6						
prétérit						1-6	
subjonctif présent		1-3, 6					4-5, (1-3, 6)
subjonctif passé						1-6	
impératif		2					5
participe						✓	
gérondif							✓
infinitif	✓						

	/al-/	/vε/	/va/	/vɔ̃/	/i-/	/aj/
présent	4-5	1	2-3	6		
futur					1-6	
conditionnel					1-6	
imparfait	1-6					
prétérit	1-6					
subjonctif	4-5					1-3, 6

<sup>158</sup> *al-e* est traditionnellement affecté au 3<sup>e</sup> groupe (classe IIa). Le mettent toutefois à part (a) l'ampleur du supplétisme des bases ; (b) le fait que, seul dans ce cas, il présente un infinitif de classe I. Son participe *al-e* 'allé.e(s)' relève aussi de cette classe, propriété qu'il partage avec *et-r* 'être' (*et-e* 'été').

<sup>159</sup> Les formes entre parenthèses sont des variantes, p.ex. (*kə*) *ʒə=swaj* pour (*kə*) *ʒə=swa* '(que) je sois'.

présent						
subjonctif passé	1-6					
impératif	5		2			
participe	✓					
gérondif	✓					
infinitif	✓					

Ces tableaux confirment l'observation du chapitre précédent, que seul le présent (de l'indicatif et, dans une moindre mesure, du subjonctif) se fléchit sur plusieurs bases.

Le verbe *fɛ-r* 'faire' requiert lui aussi un nombre important de bases : /fɛ/, /fəz-/ (cf. *nu=fəz-ɔ̃* 'nous faisons'), /fɛt/ (*vu=fɛt* 'vous faites'), /fɔ̃/ (*i(l)=fɔ̃* 'ils font'), /fə-/ (cf. *nu=fə-rɔ̃* 'nous faisons'), /f-/ (cf. *nu=f-im* 'nous fîmes'). D'autre part, on ne peut manquer de relever un certain nombre de parallèles formels entre les paradigmes d'*av-war* et d'*al-e* :

prés.ind 1	prés.ind 2-3	prés.ind 4-5	prés.ind 6	fut, cond	ipf	prés.subj
ɛ	a	av-ɔ̃	ɔ̃	o-rɛ, etc.	av-ɛ, etc.	ɛ(j)
vɛ	va	al-ɔ̃	vɔ̃	i-rɛ, etc.	al-ɛ, etc.	aj

Sont-ils le fruit du hasard ? Oui, sans doute. Ces formes que leur anomalie réunit se sont-elles néanmoins influencées mutuellement ? On ne l'exclura pas.

Comme on l'a vu, la flexion des verbes de classe I ne s'accompagne d'aucun supplétisme autre qu'apparent, en sorte qu'on peut considérer que chaque lexème se réalise sous des formes à base unique. Plusieurs bases, en revanche, sont requises pour manifester les verbes de classe IIa et IIb. Pour la suite de l'exposé, on s'en tiendra aux paradigmes des trois verbes exemplaires : *fāt-e* 'chanter' (I), *vul-war* 'vouloir' (IIa) et *ateri-r* 'atterrir' (IIb). Pour le reste, « le » Bescherelle reste l'outil indispensable, pourvu qu'on en élimine tout ce qui n'est qu'orthographe.

Les paradigmes décrits par les tableaux ci-dessous partagent une particularité : le syncrétisme (l'identité formelle) des personnes 1-3, uniquement distinguées par les clitiques personnels attachés à une forme verbale par ailleurs uniforme.<sup>160</sup> On peut estimer qu'il s'agit d'un syncrétisme motivé, explicable, puisque les trois formes partagent le trait [NOMBRE *singulier*]. Font seuls exception les trois verbes anomaux *ɛt-r* 'être', *av-war* 'avoir' et *al-e* 'aller', le futur synthétique qui, dans l'ensemble 1-3, distingue 1, et les deux paradigmes les plus marginaux, dont il est permis de questionner l'appartenance au FP : le prétérit de la classe I qui distingue 1, et le passé du subjonctif des trois classes qui distingue 3, toujours dans l'ensemble 1-3.

Les suffixes flexionnels du FP sont souvent — pas toujours — analysables en un exposant TAM suivi d'un exposant PNG : p.ex. (*vu=*)*fātə-r-e* /*fāt(ə)-r<sub>F</sub>-e<sub>S</sub>*/ 'vous chanterez'. Or, l'expression des valeurs de PNG apparaît constante pour les trois classes I, IIa et IIb et toutes les valeurs de TAM, à l'exception du futur, du prétérit et du passé du subjonctif :

1		4	ɔ̃
2		5	e
3		6	

Tableau x : suffixes PNG ailleurs qu'au futur, au prétérit et au passé du subjonctif

<sup>160</sup> A noter que *ɔ̃* 'on' reste morphologiquement singulier même quand il vaut pour *nu* 'nous'.

Aucun suffixe PNG ne correspond aux valeurs 1, 2, 3 et 6, qui ne sont exprimées que par la flexion semi-analytique en classe I :  $ʒə|ty|əl=ʃāt$  ‘je|elle chante, tu chantes, elles chantent’ ; par la flexion semi-analytique et l’alternance des bases en classe IIa et IIb :  $ʒə|ty|əl=vø$  ‘je|tu veux, elle veut’,  $əl=væl$  ‘elles veulent’ ;  $ʒə|ty|əl=fini$  ‘je|tu finis, elle finit’,  $əl=finis$  ‘elles finissent’. Le passé du subjonctif, pour autant qu’il existe encore en FP, ne fait pas exception sur ce point : dans (*kə*)  $ʒə=|ty=ʃāt-as$  ‘(que) je|tu chantasse(s), (*k*)  $əl=ʃāt-a$  ‘elle chantât’, (*k*)  $əl=ʃāt-as$  ‘elles chantassent’, les suffixes *-as* et *-a* expriment les traits de TAM, non de PNG.

Hormis les trois anomaux, font exception à la constance des suffixes PNG *di-r* ‘dire’ et *fē-r* ‘faire’, qui ont  $vu=dit$  ‘vous dites’ et  $vu=fēt$  ‘vous faites’ à 5. Plutôt que d’identifier la finale *-t* comme un suffixe alternatif — ce qu’elle fut — il me paraît plus conforme aux faits de considérer /dit/ et /fēt/ comme deux bases associées à la valeur PNG en question. Les dérivés de *di-r* sont régularisés :  $vuz=ēterdiz-e$  ‘vous interdisez’ ; mais pas ceux de *fē-r* :  $vu=defēt$  ‘vous défaites’ — quoique  $vu=defəz-e$  s’entende assez couramment.

Il suit du tableau **x** que la valeur de TAM {TEMPS *présent*, ASPECT *imperfectif*, MODE *indicatif*} s’exprime par défaut, autrement dit par l’absence d’exposant dédié, dans les verbes de classe I : cf.  $nu=ʃāt-ḥ$  ‘nous chantons’,  $vu=ʃāt-e$  ‘vous chantez’, où *-ḥ* et *-e* n’expriment que les valeurs de PNG 4 et 5. Aux autres personnes que 4-5, la base nue exprime la valeur de TAM, les clitiques personnels celles de PNG. Et de même dans les verbes des classes IIa et IIb si l’on fait abstraction des alternances de bases.

On en dirait autant de la valeur {TEMPS *présent*, ASPECT *imperfectif*, MODE *subjonctif*}, n’était le /j/ qu’on voit apparaître entre le thème et le suffixe PNG aux personnes 4 et 5. S’agit-il d’un exposant de la valeur du mode ? Sans doute pas, car le même /j/ apparaît à l’imparfait de l’indicatif aux mêmes personnes. C’est pourquoi j’analyserai plutôt les finales *-jḥ* et *-je* comme des morphes cumulatifs inanalysables, exposants des ensembles {TEMPS *présent*, ASPECT *imperfectif*, MODE *subjonctif*, PN 4} et {TEMPS *présent*, ASPECT *imperfectif*, MODE *subjonctif*, PN 5}, d’une part, {TEMPS *passé*, ASPECT *imperfectif*, MODE *indicatif*, PN 4} et {TEMPS *passé*, ASPECT *imperfectif*, MODE *indicatif*, PN 5}, d’autre part. Et j’écris (*kə*)  $nu=ʃāt-jḥ$  ‘(que) nous chantions’. L’expression des valeurs de TAM au présent du subjonctif et à l’imparfait de l’indicatif n’est donc pas entièrement par défaut, bien qu’elle ne repose pas sur un exposant qui lui soit uniquement dédié.

Les autres personnes de l’imparfait (1236) des deux classes comportent un suffixe *-ε*, exposant de la valeur de TAM {TEMPS *passé*, ASPECT *imperfectif*, MODE *indicatif*}, que je supposerai non spécifié pour les valeurs de PNG :  $ʒə|ty|əl=ʃāt-ε$  ‘je|tu chantais, elle chantait, elles chantaient’,  $ʒə|ty|əl=vul-ε$  ‘je|tu voulais, elle voulait, elles voulaient’,  $ʒə|ty|əl=finis-ε$  ‘je|tu finissais, elle finissait, elles finissaient’. Spécifiés qu’ils sont pour les valeurs 4 et 5, *-jḥ* et *-je* s’imposent dès lors que ces valeurs se voient convoquées par la flexion clitique ( $nu(z)=$  et  $vu(z)=$ ).<sup>161</sup> Le syncrétisme 1236 reproduit celui du présent (cf. tableau **x**).

En classe IIa et IIb, l’alternance des bases est liée à l’association TAM-PNG comme le montrent les tableaux ci-dessus. Mais attention : « liée à » ne veut pas dire « exprime ». La même analyse vaut pour  $əl=væl$  ‘elles veulent’ que pour  $əl=ʃāt$  ‘elles chantent’ : la base nue exprime les valeurs de TAM par défaut, c’est-à-dire du seul fait de n’être pas marquées, les clitiques personnels celles de PNG, avec pour seule différence que CHANTER n’a qu’une base, VOULOIR plusieurs. Reste qu’à exposants PNG constants, l’alternance des bases contribue à la flexion synthétique : si, dans  $ʒə|ty|əl=vø$  |  $væl$  ‘je|tu veux, elle veut, elles veulent’, le verbe se réduit à sa base, la valeur {TEMPS *présent*, ASPECT *imperfectif*, MODE *indicatif*, PNG 1236}, n’en est pas moins recouvrable, outre les clitiques personnels pour PNG, du contraste avec la base /vul/ associée à {TEMPS *présent*, ASPECT *imperfectif*, MODE *indicatif*, PN 4-5} ou

<sup>161</sup> C’est un principe général en morphologie que les exposants plus spécifiés prennent le pas sur les moins spécifiés. Panini, grammairien indien du quatrième siècle AEC, lui a attaché son nom.

{TEMPS *passé*, ASPECT *imperfectif*, MODE *indicatif*, PN 1-6} : cf. *nu=vul-ō* ‘nous voulons’, *vu=vul-e* ‘vous voulez’, *zə=vul-ε* ‘je voulais’, etc. C’est de même l’alternance des bases qui, aux personnes 123, signale seule la distinction modale entre indicatif et subjonctif présents : *zə|ty|εl=vø* ‘je|tu veux, elle veut’ vs. (*kə*) *zə|ty|εl=væj* ‘(que) je|tu veuilles, (qu’)elle veuille’. Par là, la base mutable constitue, non pas un exposant, mais un indice, certes non univoque, de la valeur de TAM.

Le futur synthétique est caractérisé par un suffixe *-rV*, dans lequel *-r-* est constant et peut donc être vu comme l’exposant de la valeur de TAM, tandis que la voyelle varie selon la valeur de PNG, dont elle semble donc constituer l’exposant : cf. *z=apresi-r-ε* ‘j’apprécierai’, *ty=fātā-r-a* ‘tu chanteras’, *nuz=ākadrā-r-ō* ‘nous encadrerons’. L’autre analyse, qui fait de *-rε*, *-ra*, etc. des suffixes cumulatifs, n’est toutefois pas déraisonnable. Pour les verbes pluribase, la base du futur est souvent la même que celle de l’infinitif : cf. *syiv-r* ‘suivre’ et *zə=syiv-r-ε* ‘je partirai’, *ātād-r* ‘entendre’ et *z=ātād-r-ε* ‘j’entendrai’. Cette correspondance, fondée sur l’histoire, connaît toutefois de nombreuses exceptions, conséquences de changements phonétiques anciens qui l’ont rendue opaque : cf. *av-war* ‘avoir’ vs. *z=o-r-ε* ‘j’aurai’, *fε-r* ‘faire’ vs. *zə=fā-r-ε* ‘je ferai’, *sav-war* ‘savoir’ vs. *zə=so-r-ε* ‘je saurai’, *val-war* ‘valoir’ vs. *zə=vod-r-ε* ‘je vaudrai’, *vwa-r* ‘voir’ vs. *zə=vε-r-ε* ‘je verrai’, *vən-i-r* ‘venir’ vs. *zə=vjēd-r-ε* ‘je viendrai’, etc. Seuls *εt-r* et *al-e* présentent au futur des bases originellement distinctes de celle de l’infinitif : *zə=sā-r-ε* ‘je serai’, *z=i-r-ε* ‘j’irai’. Bien entendu, ces considérations historiques sont sans pertinence pour notre objet. Les bases sont identiques (le plus souvent) ou différentes (pour des verbes encore assez nombreux et, qui plus est, fréquents) et il faut l’apprendre et le mémoriser, car aucune opération phonologique synchronique raisonnable ne permet de prédire, par exemple, /vjēd-/ à partir de /vən-/.

La voyelle finale exprime les valeurs de PNG, davantage distinguées qu’au présent (cf. tableau **x**). On observe deux syncrétismes, 2/3 et 4/6. Le premier participe du syncrétisme général des personnes du singulier, le second est propre au futur :

1	-ε	4	-ō
2	a	5	-e
3		6	-ō

Tableau **y** : suffixes PNG au futur synthétique

A la personne 1 je note un /ε/ ouvert, quoique l’orthoépée réclame un /e/ fermé. Mais cette prononciation n’est plus guère pratiquée, et c’est /ε/ que l’on entend le plus souvent en FV « nordiste », si bien que le futur *zə=fātā-r-ε* ‘je chanterai’ ne se distingue plus phonologiquement du conditionnel *zə=fātā-r-ε* ‘je chanterais’, tous deux /z(œ)ʃāt(œ)ʁε/.

Pour les verbes tels que *part-ir* on pourrait admettre que la voyelle finale associée aux traits PNG est simplement ajoutée à l’infinitif — ce qui est *de facto* conforme à l’origine de ces formes en latin tardif : *partire-habeo* > *partirai*. Un événement aussi ancien n’est évidemment d’aucun poids en synchronie. Surtout, cette analyse laisserait de côté non seulement les verbes de la classe I de l’infinitif desquels /r/ est absent depuis le moyen français, mais aussi ceux de la classe IIa dont les bases diffèrent à l’infinitif et au futur. Il convient donc bien d’analyser le /r/ de *parti-r-ε* comme l’exposant du futur, homophone mais distinct du /r/ de l’infinitif. Le prix à payer — mais il est modique — est d’ajouter une base liée, /parti-/ (cf. Bonami & Boyé 2003).

Le passé synthétique perfectif ou prétérit, traditionnellement nommé « passé simple », n’appartient plus à la compétence active de la plupart des francophones. Son emploi constant dans la narration écrite fait néanmoins qu’il relève toujours, jusqu’à un certain point, de leur compétence passive. Le prétérit n’existe qu’à l’indicatif. Du fait de son caractère narratif, seules les personnes 3 et 6 sont vraiment en usage. Les personnes 4 et 5 ne s’emploient plus

guère que par plaisanterie. La classe I ne présente aucune forme par défaut et pour seul syncrétisme celui des personnes 2 et 3, comme le montre le tableau ci-dessous :

1	-e	4	-am
2	a	5	-at
3		6	-εr

Tableau x : suffixes TAM-PNG au prétérit (I)

Ici, je note la désinence 1 par /e/ fermé, la distinguant ainsi de sa contrepartie /ε/ ouvert de l'imparfait (passé imperfectif). Reste que, pas davantage qu'au futur, cette opposition classique n'est maintenue dans les rares occasions où l'on articule la forme :  $ʒə=f\tilde{a}t-e$  'je chantai' se prononce en fait comme  $ʒə=f\tilde{a}t-ε$  'je chantais', soit /ʒ(œ)fātε/. Queneau s'en amuse dans ce passage de *Zazie dans le métro* (p. 156 du Livre de Poche) :

« M'autorisez-vous donc à de nouveau formuler la proposition interrogative qu'il y a quelques instants j'énonça devant vous ? — J'énonçai, dit l'obscur. — J'énonçais, dit Trouzcaillon. — J'énonçai sans esse. — J'énonçai, dit enfin Trouzcaillon. Ah ! la grammaire, c'est pas mon fort. »

La classe IIb (*ateri-r, fini-r*) est tout aussi régulière, comme le montre le tableau x :

1	(- i)	4	(-i)m
2		5	-(i)t
3		6	-(i)r

Tableau x : suffixes TAM-PNG au prétérit (IIb)

En fait, deux analyses de ce temps sont possibles, d'où les parenthèses du tableau. Selon l'une, le prétérit des verbes de la classe IIb met en jeu une base consonantique (/ater-/ , /fin-/), et c'est ce qui distingue (silencieusement)  $L=avj\tilde{s}$  *ateri* 'L'avion atterrit (en ce moment)' de  $L=avj\tilde{s}$  *ater-i* 'L'avion atterrit (hier)'. Soutient ce choix d'analyse le fait que les mêmes suffixes caractérisent le prétérit d'un certain nombre de verbes en *-ir* de la classe IIa : p.ex.  $ʒə=d\tilde{o}rm-i$  'je dormis',  $ʒə=part-i$  'je partis',  $ʒə=s\tilde{o}rt-i$  'je sortis', etc., dans lesquels il est évident que /i/ ne peut appartenir à la base. Selon l'autre analyse, prétérit et présent de la classe IIb partagent la même base vocalique, si bien que les personnes 1-3 sont homophones — on ne parlera pas de syncrétisme, notion qui ne vaut qu'au sein d'un même paradigme — et que les suffixes *-m*, *-t* et *-r* expriment cumulativement TAM et PNG 4-6. Une répugnance certaine pour les hypothèses qui mettent en jeu des structures dépourvues de support empirique, telle la « distinction » *ateri* vs. *ater-i*, me ferait préférer la seconde. Mais le débat, s'il a un intérêt, reste ouvert.

La classe IIa, en revanche, est massivement supplétive et présente une variation quant aux suffixes TAM-PNG. Pour les verbes du type de *vul-war* 'vouloir', c'est la base de l'infinitif qui apparaît à toutes les personnes, et le suffixe TAM-PNG présente un /y/, qui distingue ces verbes de ceux du type de *s\tilde{o}rt-ir* et *d\tilde{o}rmi-r*, où la base est aussi celle de l'infinitif, mais la voyelle du suffixe est /i/, dont on peut supposer qu'il se fond avec le /i/ basal (haplogie).

1	-y/-i	4	-y/im
2		5	-y/it
3		6	-y/ir

Tableau 8 : suffixes TAM-PN au prétérit IIa

Cf. *sɔrt-ir* ‘sortir’ :  $\mathfrak{z}\partial|ty|\varepsilon l=s\partial r t-i$  ‘je|tu sortis, elle sortit’,  $nu=s\partial r t-im$  ‘nous sortîmes’,  $vu=s\partial r t-it$  ‘vous sortîtes’,  $\varepsilon l=s\partial r t-ir$  ‘elles sortirent’. Comme on l’a vu, les verbes du type de *bwa-r* et *krwa-r* ‘croire’ fléchissent le prétérit sur une base monoconsonantique :  $\mathfrak{z}\partial|ty|\varepsilon l=kr-y$  ‘je|tu crus, elle crut’,  $nu=kr-ym$  ‘nous crûmes’,  $vu=kr-yt$  ‘vous crûtes’,  $\varepsilon l=kr-yr$  ‘elles crurent’. *Puv-war* ‘pouvoir’ et *sav-war* présentent aussi, exceptionnellement, une base monoconsonantique au prétérit et aux formes qui lui sont liées (passé du subjonctif et participe) :  $\mathfrak{z}\partial|ty|\varepsilon l=p-y$  ‘je|tu pus, elle put’,  $nu=p-ym$  ‘nous pûmes’,  $vu=p-yt$  ‘vous pûtes’,  $\varepsilon l=p-yr$  ‘elles purent’,  $p-y$  ‘pu.e(s)’ ;  $\mathfrak{z}\partial|ty|\varepsilon l=s-y$  ‘je|tu sus, elle sut’,  $nu=s-ym$  ‘nous sûmes’,  $vu=s-yt$  ‘vous sûtes’,  $\varepsilon l=s-yr$  ‘elles surent’,  $s-y$  ‘su.e(s)’.

Ne vaudrait-il pas mieux faire de /py/ et /sy/ des bases alternatives, en sorte que les personnes 1-3 s’exprimeraient par défaut, et que 4-6 s’analyseraient comme *py-m*, *py-t*, *py-r* et *sy-m*, *sy-t*, *sy-r* ? On est arrêté sur cette voie par l’identité de la voyelle de cette base hypothétique avec la voyelle désinentielle des autres verbes, qui pose question. Reste qu’on est contraint à cette analyse avec le prétérit d’*av-war* ‘avoir’, qui paraît réduit aux désinences :  $\mathfrak{z}|ty|\varepsilon l=y$  ‘je|tu eus, elle eut’,  $nu=z=y-m$  ‘nous eûmes’,  $vu=z=y-t$  ‘vous eûtes’,  $\varepsilon l=z=y-r$  ‘elles eurent’. Mais désinence suppose base, et, sauf à supposer une base « zéro », à quoi je me refuse, on ne peut que considérer /y/ comme la base du prétérit (et du passé du subjonctif et du participe) comme je l’ai fait au tableau **x**, si bien que les personnes 1-3 s’expriment en effet par défaut, et les consonnes désinentielles des personnes 4-6 s’attachent directement à la base.

Le paradigme du tableau **x** est aussi celui de *mudr* ‘moudre’ et de *rezudr* ‘résoudre’ :  $\mathfrak{z}\partial|ty|\varepsilon l=m\partial l-y$  ‘je|tu moulus, elle moulut’,  $nu=m\partial l-ym$  ‘nous moulûmes’, etc.,  $\mathfrak{z}\partial|ty|\varepsilon l=r\partial z\partial l-y$  ‘je|tu résolut, elle résolue’,  $nu=r\partial z\partial l-ym$  ‘nous résolûmes’, etc., le premier avec un thème /mul-/, le second /rezOl(v)-/ ailleurs qu’à l’ensemble {infinitif-futur-conditionnel}. Mais les quelques autres verbes en *-udr* sont différents : *kudr* ‘coudre’ a  $\mathfrak{z}\partial|ty|\varepsilon l=k\partial z-i$  ‘je|tu cousis’, elle cousit’,  $nu=k\partial z-im$  ‘nous cousîmes’, etc. *Absudr* ‘absoudre’ et *disudr* ‘dissoudre’ sont sans prétérit ; *surdr* ‘sourdre’ n’a que les personnes 3 et 6 du présent et de l’imparfait de l’indicatif.

L’unicité de la base est une caractéristique du prétérit, qui s’étend même aux trois anomaux :  $\mathfrak{z}\partial=f-y$  ‘je fus’,  $nu=f-ym$  ‘nous fûmes’, etc. ;  $\mathfrak{z}=y$  ‘j’eus’,  $nuz=y-m$  ‘nous eûmes’, etc. ;  $\mathfrak{z}=al-e$  ‘j’allai’,  $nuz=al-am$  ‘nous allâmes’, etc. Ce n’est toutefois qu’une aide relative en classe IIa, car cette base est souvent imprévisible et tout autant la terminaison : cf.  $nu=k\partial r-ym$  ‘nous courûmes’, mais  $nu=p\partial r t-im$  ‘nous partîmes’. On notera aussi que la relative abondance de suffixes TAM-PNG au prétérit ne suffit pas à éluder le principe général qui fait du FP une langue à sujet obligatoirement exprimé à la différence de la majorité des langues romanes : *\*kuz-im* serait parfaitement compréhensible (cf. occitan *cozeguèrem* /kuze'gɛren/), mais c’est  $nu=k\partial z-im$  ‘nous cousîmes’ qui se dit (peu).

Je n’irai pas plus loin dans l’inventaire des formes du prétérit. La flexion du FP s’y montre particulièrement opaque, sans doute la plus opaque de toutes les langues romanes. C’est là que les francophones qui se risquent à en produire à l’oral sont les plus exposés à l’erreur (voir la citation de Queneau ci-dessus). Reste que les enfants l’acquièrent, passivement — puisqu’ils le comprennent plus qu’ils ne le produisent — incomplètement, mais de façon néanmoins systématique à en croire la conclusion de l’étude de Labelle & Morris (2011: 28):

« *Two crucial points emerge from the present study. Firstly, phonological factors play an important role in the participants’ production: the phonological form of the stem, the syllabic structure of the stem and the choice of stem given the vocalic nature of the ending all explain some aspects of the children’s creative analogies. Secondly, learners take into account a variety of paradigm cells in order to compute the PS [passé simple]: the present, participle, and infinitive*



are used to compute the ending, and the imparfait is used to compute the stem. Therefore learners are sensitive not only to phonological factors, but also to paradigmatic structure. »<sup>162</sup>

Qu'imparfait et prétérit partagent leurs bases est exact pour les verbes de la classe IIa à infinitif en *-(i)r* (*part-ir*, *rãd-r* 'rendre', etc.), mais non pour la classe IIb (cf. *nu=finis-jõ* 'nous finissions' vs. *nu=fini-m* 'nous finîmes') ni, on vient de le voir, pour les IIa à infinitif en *-war*, de tous les plus « irréguliers ». Cela étant, les observations de Labelle & Morris, faites sur de jeunes Québécois.e.s, mais certainement extensibles à l'ensemble des francophones natifs, confirment au mieux l'importance des espaces thématiques, des paradigmes et de l'analogie, « créative » en cela qu'elle permet d'inférer (*compute*) des formes encore jamais entendues, qui se révéleront justes ou fausses, mais apparemment assez souvent justes. Il n'en est pas moins assez probable que les formes les plus « déviantes » ( $\mathcal{J}=y$ ,  $\mathcal{J}\partial=b-y$ ,  $\mathcal{J}\partial=p-y$ ,  $\mathcal{J}\partial=s-y$ , etc.) soient mémorisées en bloc et à l'unité.

Le subjonctif passé est étroitement associé au prétérit pour autant que l'un et l'autre partagent toujours le même thème. Il n'appartient sans doute pas au FP, et plus guère au FO. Mais il est toujours possible de l'introduire dans un discours ordinaire, le plus souvent par plaisanterie, à divers degrés d'exactitude. La forme est en sommeil, elle n'est pas morte. Presque personne ne la connaît bien, mais beaucoup en ont des bribes. C'est pourquoi on ne peut l'ignorer.

A la classe I les terminaisons sont comme indiqué au tableau **x**.

1		4	-asjõ
2	-as	5	-asje
3	-a	6	-as

Tableau **x** : suffixes TAM-PNG au subjonctif passé (I)

On reconnaît en 4 et 5 les morphes /jõ/ et /je/ identifiés aux mêmes personnes du subjonctif présent et de l'imparfait. D'où une possible décomposition en deux parties : /as/ exposant de {TEMPS *passé*, ASPECT *imparfaitif*, MODE *subjonctif*, PNG 1-2, 4-6} et /jõ/ et /je/ exposants de 4 et 5 respectivement. Que les mêmes traits PNG soient alors exprimés deux fois n'est pas en soi un problème ; cela porte un nom technique, « exponence multiple », et s'observe assez fréquemment dans les langues naturelles. Il me paraît toutefois plus simple et sans doute plus conforme à ce qui peut subsister de sentiment linguistique en la matière de considérer tous les suffixes du tableau **x** comme cumulatifs et inanalysables.

Les mêmes se retrouvent à la classe IIb, mais la voyelle est /i/. Si, en classe I, j'ai rattaché à la terminaison la voyelle /a/, c'est que celle-ci, quoique traditionnellement considérée comme thématique, n'appartient pas de fait aux bases des verbes de cette classe. (Seule l'invocation de l'étymon *cantare* /kant-a-re/ de *řãt-e* permettrait de soutenir le contraire.) Il en va différemment en classe IIb. La voyelle y est celle du thème court : (*kə*)  $\mathcal{J}\partial=$  | *ty=fini-s* '(que) je | tu finisse(s)', (*k=*)*el=fini* '(qu')elle finît, (*kə*) *nu=fini-sjõ* '(que) nous finissions', etc. Il s'ensuit l'identité formelle des subjonctifs présent et passé aux personnes en /s/. Mais, comme on le voit, une identité qui recouvre deux analyses distinctes : thème long /finis/ au présent, nu aux personnes 1-3 et 6 ; thème court /fini/ au passé, suivi de l'exposant /s/ à toutes les personnes sauf 3.

<sup>162</sup> « Deux points essentiels ressortent de la présente étude. Premièrement, les facteurs phonologiques jouent un rôle important dans les productions des participant.e.s: la forme phonologique et la structure syllabique du thème, ainsi que son choix du fait de la nature vocalique de la terminaison, tout cela explique certains aspects des analogies créatives des enfants. Deuxièmement, les apprenant.e.s prennent en compte plusieurs cases du paradigme afin d'inférer la forme du PS [passé simple]: le présent, le participe et l'infinitif servent à inférer la terminaison, l'imparfait sert à inférer le thème. Les apprenant.e.s sont donc sensibles non seulement à des facteurs phonologiques, mais aussi à la structure paradigmatique. »

C'est ce même exposant associé à la voyelle /y/ ou /i/ qui, dans les verbes de la classe IIa, s'attache à la base de prétérit : (*kə*) *ʒə=b-ys* '(que) je busse', (*kə*) *vu=b-ysje* '(que) vous bussiez', etc. ; (*kə*) *ʒə=part-is* '(que) je partisse', (*kə*) *vu=part-isje* '(que) vous partissiez', etc.

La matérialité phonique de cet exposant mérite par ailleurs quelques commentaires. A la différence du prétérit, le subjonctif passé est banni de l'oral, sauf plaisanterie, et pratiquement inutilisable à l'écrit, à nouveau sauf intention comique, car il s'y attache un ridicule qui n'entache pas le prétérit. Pourquoi un tel opprobre sur des formes que les autres langues romanes emploient communément ? Pourquoi ne dit-on plus comme encore au dix-huitième siècle *Il=falε kə ty=fāt-as* 'Il fallait que tu chantasses' quand le portugais, par exemple, dit encore *Precisava que cantasses* /prəsi'zavə kə\_kã<sup>n</sup>'tasəf/ ? Le système de la concordance des temps n'est pas en cause, qui fonctionne très bien à l'indicatif. Reste la phonétique et la sensibilité des francophones aux calembours : comment résister au jeu de mots salace que suscite *Il=vul-ε kə ty=lə=/sys/* 'Il voulait que tu le susses (*s-ys*) / suces (*sys*)' ? A quoi s'ajoutent les connotations péjoratives de la terminaison /as/ : *degolas* 'dégueulasse' — « Qu'est-ce que c'est... ? » — *petas* 'pétasse', *popylas* 'populace', *salas* 'salace', etc. Que ces fréquentations douteuses aient fortement contribué au rejet du subjonctif passé des verbes de la classe I, d'un nombre important de verbes de la classe IIa et des trois anomaux, voire l'aient causé, le confirme le fait que la seule forme que l'on s'autorise encore assez librement à l'écrit et parfois à l'oral est la personne 3, la seule sans /s/. La classe IIb, quoique épargnée par ces malsonnances, n'a pu que suivre le mouvement. Précisons encore que l'unique forme entendue avec quelque fréquence en FP — et dans les registres les plus surveillés de celui-ci — est *fy* 'fût', qui a pour elle la discrétion : *Il fale k=il=fy(t)\_idjo* 'Il fallait qu'il fût idiot'.

## 10.2. Les formes finies analytiques

Comme expliqué au chapitre 5, la flexion analytique consiste en la collocation de deux verbes, l'un dit « auxiliaire », l'autre « principal », le tout formant une périphrase et appartenant au paradigme du verbe principal à l'instar des formes synthétiques. Une analyse adéquate du FP exige toutefois de distinguer deux sortes de périphrases verbales : les temps-aspects analytiques et les périphrases aspecto-temporelles. Les premiers comprennent le prétérit analytique (« passé composé »), le passé antérieur (« plus-que-parfait ») et le futur antérieur. Ce sont eux qu'Abeillé & Godard (2021, chapitre 3) disent « fusionnés », car leurs formes, bien que faites de plusieurs éléments modérément séparables, se comportent comme un seul. Les secondes sont le futur analytique, le prétérit immédiat et le progressif (ou duratif). On peut ainsi distinguer entre auxiliaires tout court et auxiliaires aspectuels.<sup>163</sup>

En FP, y compris dans ses variantes les plus proches du FO, le prétérit analytique s'est substitué au synthétique dans l'usage commun. L'auxiliaire est *av-war* 'avoir', à moins que le verbe principal ne soit inaccusatif (*al-e* 'aller', *mōt-e* 'monter', *part-ir* 'partir', etc.) ou réfléchi, auquel cas *εt-r* 'être' est l'auxiliaire.<sup>164</sup> Description bien sûr trop sommaire et qui ne tient pas compte des variations possibles où la grammaire prescriptive voit des fautes.<sup>165</sup> Mais nous pouvons nous en contenter. Le verbe principal est à la forme participiale (voir chapitre 11). Dès lors donc qu'on sait fléchir *av-war* et *εt-r* et qu'on connaît les participes — là gît la difficulté, car ils sont souvent peu prévisibles — on maîtrise le prétérit analytique, ainsi que le

<sup>163</sup> On trouve aussi dans les grammaires « semi-auxiliaires » pour désigner les auxiliaires aspectuels tels qu'ALLER du futur analytique, sans doute afin d'insister sur le fait qu'ils fonctionnent aussi comme des verbes pleins. Mais il en va de même d'ÊTRE et AVOIR (*Je pense donc je suis, J'aurai de l'or*). Ce semble être une propriété très générale des auxiliaires, qui connaît peu de contre-exemples, qu'ils soient aussi des verbes pleins.

<sup>164</sup> On qualifie d'inaccusatifs les verbes *intransitifs* dont le sujet est affecté par l'événement plutôt qu'il ne le cause. *Tōbe* 'tomber' en est un exemple évident.

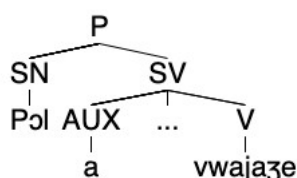
<sup>165</sup> Lorsqu'un verbe dispose de plusieurs structures argumentales, l'auxiliaire peut varier en conséquence : cf. *Il=ε tōb-e* 'Il est tombé' et *Il=a tōb-e la=vest* 'Il a tombé la veste'. Pour certains verbes, la variation est passée dans l'usage : cf. *Sə=livr ε/a par-y l=ā dernje* 'Ce livre est/a paru l'an dernier'.

passé et le futur antérieurs qui n'en diffèrent que par le temps de l'auxiliaire. Comme on l'a vu, auxiliaire et verbe principal peuvent être séparés par un adverbe et/ou une incise : *il=a, zə=krwa, boku\_pl-y l=ete\_dernje* 'Il a, je crois, beaucoup plu l'été dernier', *zə=mə=suji\_teriblāmā\_ānyij-e* 'Je me suis terriblement ennuyé' — conséquence de la double nature des périphrases flexionnelles : mots-formes et syntagmes. Plus précisément, cette duplicité signifie que, par exemple, *pəl a, zə=krwa, boku\_vwaja3-e* 'Paul a, je crois, beaucoup voyagé' appelle deux analyses parallèles. En tant que mot-forme, membre du paradigme du lexème VOYAGER, *a vwaja3-e* peut se représenter comme suit :

[VOYAGER : prétérit analytique, 3SG] ↔ [a<sub>AUXAVOIR, 3SG...</sub> vwaja3-e<sub>PTCP</sub>]

En clair : à la 3<sup>e</sup> personne du singulier du prétérit analytique du verbe VOYAGER correspond la collocation discontinue (...) de la 3<sup>e</sup> personne du singulier de l'auxiliaire AVOIR et du participe du verbe, et inversement.

En tant que syntagme, la même forme est justiciable d'une représentation arborescente comme ci-dessous :



Ces deux représentations du même objet devront bien sûr être complétées d'une caractérisation des éléments susceptibles d'occuper la position des trois points.

Le futur analytique se forme au moyen de la collocation ALLER + V infinitif. Les conditions qui en commandent l'usage plutôt que du futur synthétique sont complexes et n'entrent pas dans le champ du présent chapitre (mais voir chapitre 13). Y entre en revanche le fait que l'auxiliaire aspectuel n'a pas le même empan flexionnel que le verbe plein correspondant, limité qu'il est au présent de l'indicatif (*zə=vε\_(bjēto)\_fē-r la=vesel* 'je vais (bientôt) faire la vaisselle') et à l'imparfait de l'indicatif (*z=al-ε\_(zystāmā)\_fē-r la=vesel kã...* 'j'allais (justement) faire la vaisselle quand...'). A toutes les autres valeurs de TAM, on a affaire au verbe plein : cf. *zə=suji(z)=al-e\_fē-r la=vesel* 'je suis allé faire la vaisselle', *z=i-re\_fē-r la=vesel* 'j'irai faire la vaisselle', *il=fo kə z=aj\_fē-r la=vesel* 'il faut que j'aille faire la vaisselle', *zə=dwa\_al-e\_fē-r la=vesel* 'je dois aller faire la vaisselle', etc. Cette sorte de différence entre auxiliaire (aspectuel et non) et verbe plein correspondant est assez commune : en catalan, *anar* 'aller' n'est auxiliaire du prétérit analytique qu'au présent de l'indicatif ou du subjonctif : cf. *vam esmorzar* 'nous déjeunâmes, nous avons déjeuné' vs. *anirem esmorzar* 'nous irons déjeuner' (Fabra 1941, chap. 7).<sup>166</sup> En outre, le paradigme du présent de *anar* diffère aux personnes 4 et 5 selon qu'on a affaire à l'auxiliaire ou au verbe plein : cf. *vam esmorzar* 'nous déjeunâmes, nous avons déjeuné', *vau esmorzar* 'vous déjeunâtes, vous avez déjeuné' vs. *anem esmorzar* 'nous allons déjeuner', *aneu esmorzar* 'vous allez déjeuner', où *anar* retrouve son sens plein de verbe de mouvement. (Le futur analytique serait *anem a esmorzar*, et il s'emploie peu car aisément confondu avec le prétérit analytique.)

Le prétérit immédiat met en jeu l'auxiliaire aspectuel VENIR\_DE + V infinitif. Il connaît la même limitation de ses ressources flexionnelles que le futur analytique, sauf pour la possibilité du futur : cf. *zə=vjē\_(a=l-ēstā)\_də=fē-r la=vesel* 'je viens (à l'instant) de faire la vaisselle', *zə=vən-ε\_də=fē-r la=vesel kã...* 'je venais de faire la vaisselle', *kã t(y)=arivə-ra*

<sup>166</sup> Pour déroutante qu'elle paraisse, la périphrase catalane n'est pas sans équivalent en français, je veux parler de l'emploi très fréquent du futur analytique pour narrer des suites d'événements passés : p.ex. « Alors Verlaine va quitter sa femme, Rimbaud et lui vont fuir à Londres, etc., etc. »

$zə=vj\tilde{e}d-re\ də=f\tilde{e}-r\ la=vesel$  ‘quand tu arriveras, je viendrai de faire la vaisselle’. Mais c’est que l’auxiliaire VENIR\_DE, à la différence d’ALLER, est analytiquement distinct du verbe plein VENIR : cf.  $zə=vj\tilde{e}\ də=f\tilde{e}-r\ la=vesel$  ‘je viens de faire la vaisselle’ vs.  $zə=vj\tilde{e}\ də=m\tilde{o}tr\ae j$  ‘je viens de Montreuil’. Avec le verbe plein,  $də$  ‘de’ est une préposition locative ; dans la périphrase, on y verra plutôt une particule de liaison (*linker*), qui n’est autre chose que la préposition vidée de son sens, entièrement grammaticalisée.

Enfin, la périphrase progressive (ou durative) consiste en la collocation de l’expression complexe ÊTRE\_  $\tilde{a}tr\tilde{e}d\tilde{a}=\tilde{a}$  et d’un verbe infinitif. J’écris  $\tilde{a}tr\tilde{e}d\tilde{a}$  en un seul mot, car, sauf influence de l’orthographe, la forme ne paraît guère décomposable.<sup>167</sup> Quel rapport entre la syllabe /trē/ qu’elle contient et un train (sur rails ou des équipages) ? Même le « train » de « aller son train », plus ou moins à l’origine de la forme, s’en trouve à présent bien éloigné. La définition traditionnelle de locution prépositionnelle ou de préposition complexe lui convient bien, mais une préposition qui, comme le  $də$  ‘de’ du prétérit immédiat, s’est vidée de tout sens propre et n’a plus de signification que comme composante de la collocation. A la différence des précédentes, la périphrase progressive se fléchit pour toutes les valeurs TAM — p.ex.  $zə=sə-re\ (presizem\tilde{a})\ \tilde{a}tr\tilde{e}d\tilde{a}=f\tilde{e}-r\ la=vesel$  ‘je serai (précisément) en train de faire la vaisselle’,  $z=\varepsilon\ suv\tilde{a}\ et-e\ \tilde{a}tr\tilde{e}d\tilde{a}=f\tilde{e}-r\ la=vesel\ k\tilde{a}\ ty=v\tilde{a}n-\varepsilon\ m\tilde{a}=vwa-r$  ‘j’ai souvent été en train de faire la vaisselle quand tu venais me voir’. (Comme le montre ce dernier exemple, l’emploi du prétérit requiert des contextes particuliers.)

## 11. Les formes non-finies du verbe

Le FP en compte trois : l’infinitif, le participe et le gérondif. Ce dernier est généralement inconnu des grammaires scolaires, qui emploient le terme de « participe présent ». D’où une confusion entre une forme indiscutablement verbale, comme dans  $(\tilde{a}=)fasin\tilde{a}\ l=as\tilde{a}ble, \varepsilon l...$  ‘(en) fascinant l’assemblée, elle...’, et une forme adjectivale, comme dans  $yn=fam\ fasin\tilde{a}t$  ‘une femme fascinante’, qui se distingue de la première du fait de se fléchir en genre.<sup>168</sup> C’est pourquoi, soucieux d’éviter cette confusion, je choisis de nommer la forme verbale, la seule qui m’importe ici, « gérondif ». Sa marque caractéristique est le suffixe  $-\tilde{a}$  attaché au thème unique des verbes du groupe I ( $f\tilde{a}t-\tilde{a}$  ‘chantant’), au thème en /s/ des verbes du groupe IIb ( $f\tilde{a}n\tilde{a}-\tilde{a}$  ‘finissant’) et à l’un des thèmes des verbes du groupe IIa, le plus souvent celui de l’imparfait ( $al\tilde{a}$  ‘allant’,  $byv-\tilde{a}$  ‘buvant’,  $kren-\tilde{a}$  ‘craignant’,  $saf-\tilde{a}$  ‘sachant’, etc.).<sup>169</sup> Quoique invariable, le gérondif est, comme les adjectifs homophones, pourvu d’une consonne « latente » qui peut donner lieu à liaison en FP surveillé ou hypercorrect : cf.  $\tilde{a}n=a-l\tilde{a}(t)\_o=byro, z=\varepsilon...$  ‘En allant au bureau, j’ai...’. Employé seul, il sonne plutôt littéraire : cf.  $al-\tilde{a}(t)\_o=byro, z=\varepsilon...$  ‘Allant au bureau, j’ai...’. La préposition  $\tilde{a}(n)$  ‘en’ le précède le plus souvent, elle est même quasi obligatoire dans certaines expressions, moyennant une certaine influence de la position : cf.  $zə=m\tilde{a}=s\tilde{y}i\ f\tilde{e}\ mal\ *(\tilde{a})\ t\tilde{s}b-\tilde{a}$  ‘Je me suis fait mal \*(en) tombant’ vs.  $(\tilde{a})\ t\tilde{s}b-\tilde{a}\ zə=m\tilde{a}=s\tilde{y}i\ f\tilde{e}\ mal$  ‘(En) tombant je me suis fait mal’. Son absence est normale, en revanche, avec  $\varepsilon tr$  ‘être’ et  $avwar$  ‘avoir’ : cf.  $\varepsilon t-\tilde{a}(t)\ \tilde{a}=b\tilde{o}n\ s\tilde{a}te, z=estim\ k\tilde{a}...$  ‘Etant en bonne santé, j’estime que...’,  $\varepsilon j-\tilde{a}(t)\ yn=b\tilde{o}n\ s\tilde{a}te, z=estim\ k\tilde{a}...$  ‘Ayant une bonne santé, j’estime que...’.

<sup>167</sup> En tayo, créole français parlé en Nouvelle-Calédonie,  $\tilde{a}tr\tilde{e}d\tilde{a}$  a donné  $atra^nd\tilde{e}$  (/<sup>n</sup>d/ une consonne prénasalisée), marqueur de l’aspect progressif : cf.  $Ta\ atra^nd\tilde{e}\ fe\ kwa\ ?$  {tu PROG faire quoi} ‘Qu’est-ce que tu fais ?’ (cf. Kihm 1995 ; Ehrhart & Revis 2013).

<sup>168</sup> Il est vrai que la langue a longtemps confondu les deux. Le participe présent verbal (alias gérondif) pouvait s’accorder au dix-septième siècle : « Les morts se ranimants à la voix d’Elisée » (Spillebout 1985 : 396, citant Racine). Mais cet accord était déjà inaudible, sauf au féminin : « ... la veuve d’Hector pleurante à vos genoux » (Spillebout 1985 : 397, citant Racine).

<sup>169</sup> Le gérondif  $saf\tilde{a}$  est exceptionnel en cela qu’il se forme sur le thème du subjonctif présent et qu’il diffère formellement de l’adjectif-nom  $sav\tilde{a}(t)$  ‘savant(e)’.

Il y a peu à dire de l’infinitif du point de vue morphologique. Ses désinences ont été indiquées. J’ajouterai seulement que, même lorsqu’il paraît substantivé et remplir la fonction grammaticale d’un SN (p.ex. *fym-e ε\_move pur=la=sāte* ‘Fumer est mauvais pour la santé’), l’infinitif français (FP et FO) reste verbal, comme le montre l’impossibilité de lui adjoindre des déterminants (\**lā=fume ε...*). Une expression portugaise telle que *um estalar de risadas* (Teyssier 1976 : 232), littéralement « un éclater de rires », ne peut ainsi se traduire que par « des éclats de rire ».<sup>170</sup>

La forme du participe est entièrement prévisible pour les verbes des classes I et IIb : *-e* suffixé à la base unique pour les premiers : *fāt-e* ‘chanté.e(s)’ ; la base courte pour les seconds, dont le participe est l’une des significations par défaut : *fini* ‘fini.e(s)’. (Supposer un /i/ suffixé à une base consonantique /fin-/ paraît artificiel, car celle-ci n’apparaîtrait nulle part ailleurs, alors que /fini/ est assuré. Quant à supposer une haplologie /fini-i/ > /fini/, pourquoi pas ? Le défaut paraît plus simple.) Le groupe IIa se signale, comme d’habitude, par davantage de diversité. Comme dit plus haut, *al-e* ‘aller’ et *et-r* ont leur participe en classe I : *al-e* ‘allé.e(s)’, *et-e* ‘été’. Certains verbes à infinitif en *-ir* présentent le participe en *-i* que je viens d’évoquer avec réticence pour les verbes IIb, attaché au thème de l’infinitif : p.ex. *dorm-i* ‘dormi’, *fū-i* ‘fui’, *kæj-i* ‘cueilli.e(s)’, *part-i* ‘parti.e(s)’, *serv-i* ‘servi.e(s)’, *sōrt-i* ‘sorti.e(s)’, etc. D’autres ont un participe en *-y*, p.ex. *tān-y* ‘tenu.e(s)’, *vān-y* ‘venu.e(s)’, etc. D’autres encore sont pourvus d’une forme longue (alias consonne « latente »), p.ex. *ekr-i(t)* ‘écrit.e(s)’, *uver(t)* ‘ouvert.e(s)’. On se dispensera ici d’un inventaire plus complet, qu’on trouvera dans les dictionnaires et chez Bescherelle. Le moins qu’on puisse dire est que la formation des participes de la classe IIa est particulièrement opaque.

Quant à la flexion, les participes sont des adjectifs, dont ils ont d’ailleurs l’emploi à côté de leur participation aux périphrases temporelles : *le=z=ilyzjō\_perd-y* ‘les illusions perdues’, *lā=tā\_rātruv-e* ‘le temps retrouvé’, *la=mer\_al-e\_avek\_lā=solej* ‘la mer allée avec le soleil’. Comme la plupart des adjectifs, ils sont invariables en nombre ; comme eux, ils peuvent se fléchir pour le genre grammatical. Ce qui soulève la redoutable question de l’accord du participe. Elle s’aborde sous deux angles : syntaxique, pour quoi je renvoie à Abeillé & Godard (2021, chapitre 3) ; et morphologique, dont je vais traiter.

En FP, la question ne se pose pas pour les classes I et IIb dont les participes sont complètement invariables : cf. *le=poezi kā z=ε\_em-e* ‘les poésies que j’ai aimées’, *el=sā=sō(t)\_asup-i* ‘Elles se sont assoupies’. Ni pour les participes IIa sans forme longue : *le=sāriz kā z=ε\_kæj-i* ‘les cerises que j’ai cueillies’, *el=sō\_vān-y* ‘Elles sont venues’. Elle se pose, en revanche, pour les autres, susceptibles de se fléchir en genre, c.-à-d. d’exprimer le féminin au moyen de la forme longue : cf. *le=poezi kā zā=n=ε=pa(z)\_ekr-i(t)* ‘les poésies que je n’ai pas écrites’, *el=sā=sō(t)\_uver(t)* ‘Elles se sont ouvertes’.

On connaît la règle du FO : le participe composé avec l’auxiliaire *être* s’accorde avec le sujet ; composé avec *avoir*, il s’accorde avec l’objet direct si celui-ci le précède, autrement il ne varie pas. On sait aussi qu’elle est en partie artificielle. En ancien français, l’accord reflétait assez fidèlement les relations sémantiques. On accordait avec le sujet dans *Elles sont tombées* /tō<sup>m</sup>beəs/ parce que le participe le qualifie (« Elles ont la propriété d’être tombées ») ; on se dispensait d’accorder dans la composition en *avoir*, ou alors avec l’objet direct conformément à la signification résultative originelle (« avoir X en l’état d’être Yé »), et cela quelle que soit la position de celui-ci : cf. *et voit la nef autretel comme il l’avoit ore veüe* ‘et il voit la nef autrement qu’il l’avait vue alors’, *De cel remanant que Diex te lessa as tu fetes les granz proescs par les estranges terres* ‘Du restant que Dieu te laissa tu as fait les

<sup>170</sup> *Lā=rir* ‘le rire’ est à l’évidence purement nominal, comme le montre l’impossibilité de \**lā=rir o=z=ekla* ‘le rire aux éclats’. Il n’en fut pas toujours ainsi. Gougenheim (1951 : 137) cite, de Montaigne, *En certain abbayer [aboyer] du chien, le cheval cognoist qu’il y a de la colere*. Aragon ose encore « la douleur du partir », mais il est permis d’y voir quelque affectation.

grandes prouesses par les terres étrangères' (Moignet 1988 : 205-207). La règle actuelle, pronée mais encore instable en moyen français et en français classique (cf. Gougenheim 1951 : 226-227 ; Spillebout 1985 : 398-402), rigidifie ces relations que la grammaticalisation des formes, désormais uniquement temporelles, a de toute façon vidées de leur contenu.

Quoi qu'il en soit, la règle ne peut s'appliquer en FP qu'aux participes « irréguliers » bifformes.<sup>171</sup> Le fait est qu'elle s'applique : on entend *La=letr kə z=ε\_ekrit* 'La lettre que j'ai écrite'. Mais au moins autant *La=letr kə z=ε\_ekri* 'La lettre que j'ai écrite'. Il y a là une grande variation, complexe et dont je ne prétends pas comprendre les principes.

Le participe des verbes transitifs est intrinsèquement passif (plutôt que « passé ») : cf. *la=vil detruyt (par=ē=seism)* 'la ville détruite (par un séisme)', d'où son emploi en composition avec *et-r* 'être' pour former la voix passive. Comme il s'agit d'une formation entièrement analytique, je n'ai rien de plus à en dire. Le participe des verbes intransitifs, qui n'ont pas de passif, est actif comme superbement illustré par l'exemple rimbaldien de « la mer allée... »

Enfin, l'homophonie du participe et de l'infinitif pour les verbes de la classe I (*lave* 'laver|lavé.e(s)') mérite un commentaire. Elle est, on le sait, la source de fréquentes « fautes d'orthographe » dans les textes non ou mal relus produits par toute sorte de personnes, y compris de fort honnêtes, où on lit des choses comme « Elle a appeler hier » ou « Je vais rappelé aujourd'hui ». Voir aussi le fameux et déjà cité « Omar m'a tuer » qui plongea les enquêteurs dans la perplexité quant à l'identité de l'auteur du graffiti. Ou encore cet avis collé sur la porte d'un restaurant : « Fermer ce soir » (observation personnelle). Je le répète : il va de soi que, s'il avait fallu rassurer la clientèle, l'avis aurait porté « Ouvert ce soir » et non « Ouvrir ce soir ». Et personne, que je sache, ne dit ni n'écrit *\*Il=ε\_partir ijer* 'Il est partir hier' ni *\*Il=va\_parti ozurduji* 'Il va parti aujourd'hui'. Lorsque les deux formes ne se confondent pas pour la phonologie, on ne les intervertit jamais. Le problème n'est donc pas de grammaire, mais entièrement d'orthographe. Ce qui est atteint, par ignorance ou inattention, c'est la relation entre les formes et leurs représentations graphiques, relation qui, pour des raisons historiques, est devenue problématique au dix-septième siècle quand le /r/ final des infinitifs, amui en moyen français mais conservé à l'écrit, s'est vu rétabli dans la prononciation, sauf pour les verbes de la classe I (cf. Zink 1990 : 15).

## 12. La polarité

La polarité est la propriété pour une expression d'affirmer ou de nier la réalité d'un état de fait : *I(l)=plø* 'Il pleut' vs. *I(l)=(nə)=plø=pa* 'Il ne pleut pas'. En général, la polarité positive (affirmation) représente le cas par défaut et n'est pas marquée.<sup>172</sup> Que le marquage de la polarité négative relève souvent de la flexion, l'examen des langues du monde le démontre aisément (cf. Creissels 2006b, chap. 29). L'exemple du wolof suffira : cf. *dem naa* /'dɛm-na:/ {partir-PRF.1SG} 'je suis parti.e' vs. *demuma* /'dɛm-u-mʌ/ {partir-NEG-1SG} 'je ne suis pas parti.e'.

En FP, la polarité négative se marque sur le mode semi-analytique et s'exprime par les négateurs enclitiques *pa(s)* (négation absolue) ou *ply(s)* (passage du positif au négatif), variablement redoublés du négateur proclitique *nə* devant le premier clitique objet ou, à défaut, le verbe : *vu=(nə)=lə=lwi=fāt-e=pa|ply* 'Vous (ne) le lui chantez pas|plus'.<sup>173</sup> Au prétérit analytique, *(nə)...* *pa|ply* encadre l'auxiliaire, éventuellement précédé de clitiques

<sup>171</sup> Il n'existe pas, me semble-t-il, de contextes permettant l'insertion de *z* signalant la pluralité du participe. La liaison est en principe exclue dans, par exemple, *les livres que j'ai lus autrefois* (*\*/lyzotrəfwa/*). Et même si elle se fait à l'occasion, par hypercorrection, l'invariabilité flexionnelle du participe n'en sera pas mise en cause vu l'analyse de *z* ici défendue.

<sup>172</sup> La particule affirmative *que* /ke/ de l'occitan gascon est une exception bien connue : p.ex. *La gojata que drom sus ua cadièra* 'La jeune fille dort sur une chaise' (Leclerc & Javaloyès 2004 : 14).

objets :  $vu=(nə)=lə=lwi=av-e=pa|ply\_fāt-e$  ‘Vous (ne) le lui avez pas|plus chanté’. A l’infinitif, *pa* et *ply* sont proclitiques, éventuellement précédés de *nə* :  $dā=sə=pei\ i(l)=vo\_mjø\ (nə)=pa=bwa-r\ l=o\_dy=robine$  ‘Dans ce pays, il vaut mieux (ne) pas boire l’eau du robinet’,  $(nə)=pa=bwa-r\ syfizamã\ se\ move\ pur=la=sāte$  ‘(Ne) Pas boire suffisamment, c’est mauvais pour la santé’.

Comme indiqué, *pa(s)* et *ply(s)* sont biformes. Si le successeur est à initiale vocalique, il est possible de faire entendre la forme longue dont le /s/ final se voise automatiquement : cf.  $z(ə)=(n)=ε=pa(s)|ply(s)=y\_l=okazjō\ də...$  ‘Je n’ai pas|plus eu l’occasion de...’<sup>174</sup> Mais l’hiatus — /pay|ply/ au lieu de /pazy|plyzy/ — est ordinaire. D’une façon générale, il est bon de le répéter, le FP ne craint pas les hiatus, semblable en cela à l’ancien français.

Selon une analyse traditionnelle, historiquement fondée et adaptée au FO, *nə* ‘ne’ représente la négation fondamentale (le « discordantiel »), que le « forclusif » post-verbal vient renforcer (Damourette & Pichon 1911-1927, chap. 7). Voilà qui ne convient guère au FP, où c’est sûrement *pa(s)|ply(s)* qui est la négation fondamentale. C’est pourquoi, aux dénominations pittoresques des deux grammairiens-psychanalystes, je préfère les termes plus descriptifs de négateur proclitique et négateur enclitique, qui ont l’avantage de faire entrer de plein pied la polarité négative dans la flexion semi-analytique.<sup>175</sup>

Une opinion répandue veut que le négateur proclitique ait disparu du FP. Milner (1979 : 80) s’en fait l’écho : « Je ne tiendrai compte que du français écrit et ne discuterai pas les usages dits populaires, où la négation ne comporte pas *ne*. » Rien ne justifie une affirmation aussi catégorique. Il suffit d’écouter pour constater la fréquence des énoncés « à double négation », même en discours familier. L’influence du FO est probable, mais elle n’explique pas tout. J’émettrais volontiers l’hypothèse que  $zə=sε=pa|ply$  ‘Je sais pas|plus’ et  $zə=nə=sε=pa|ply$  ‘Je ne sais pas|plus’ appartiennent tous deux à toutes ou presque toutes les variétés du FP et ne sont pas pragmatiquement équivalents. Déterminer les conditions de leurs usages respectifs devra faire l’objet d’une enquête. Quoi qu’il en soit, l’omission de *nə* n’est en aucune façon récente, pas plus qu’elle ne fut toujours cantonnée à l’oral : « C’est un phénomène très courant dans la première partie du siècle [dix-septième], un peu moins fréquent par la suite » (Spillebout 1985 : 368 ; voir aussi Gougenheim 1951 : 217-220 pour le seizième siècle). Elle était plus fréquente dans les phrases interrogatives — « Peut-il pas languir à son aise / En la prison de vos cheveux ? » (Spillebout 1985 : 369 citant Malherbe) — mais ne s’y limitait pas : « (Mes affaires) elles me touchent pas tant que le malheur qui vient d’arriver au surintendant » (Spillebout 1985 : 369 citant une lettre de La Fontaine).

### 13. Français parlé et français orthographique : similitudes et différences

La description de la flexion du FP est maintenant achevée, au moins pour l’essentiel. Elle s’est ouverte sur une prise de position assez radicale, à savoir qu’un fossé sépare FP et FO dans ce domaine. Il convient à présent d’être plus précis quant à la nature de cette séparation,

<sup>173</sup> La prononciation /py/, stigmatisée, me paraît moins répandue que ne l’affirment certains observateurs du français « populaire ». Il n’est évidemment question ici que de la négation simple symbolisée par ¬P en notation logique. Dans les négations modalisées, *zame* ‘jamais’, *persøn* ‘personne’, *rjē* ‘rien’, etc. remplacent *pa* et *ply*. Je ne dirai rien du caractère « semi-négatif » ou « à polarité négative » de ces items, sur lequel FP et FO diffèrent, mais qui n’est pas de mon ressort. On ne citera que pour mémoire ‘point’, ‘mie’ et ‘goutte’, qui relèvent du FO, sont encore compris, mais sont sortis de l’usage, surtout les deux derniers, à l’exception, bien sûr, de l’expression figée « n’y voir goutte ».

<sup>174</sup> On pourrait aussi considérer que la consonne finale de la forme longue est /z/. C’est par respect pour l’histoire que je garde /s/. Et aussi parce que la consonne finale de *plys* ‘plus’ voulant dire « davantage » est /s/, toujours prononcé à la pause : cf.  $z=ā=vø\ plys$  ‘J’en veux plus’.

<sup>175</sup> Cela étant, il faut reconnaître que « discordantiel » est bien trouvé, tant il est difficile de rencontrer des énoncés négatifs qui n’expriment pas en même temps, voire d’abord un désaccord avec quelque interlocuteur réel ou virtuel.

de déterminer en quoi les systèmes flexionnels du FP et du FO diffèrent et en quoi ils sont semblables.

Comme expliqué au chapitre 2, la relation entre FP et FO est particulière, diglossie d'un genre apparemment peu commun, telle que les deux langues en présence sont assez unanimement considérées, contre toute évidence, comme deux exécutions de la même (*le français*) — au sens où deux exécutions de la Neuvième Symphonie sont toujours la Neuvième Symphonie — l'une exclusivement écrite, l'autre seulement orale, alors même que le système grammatical de l'oral, sa flexion en particulier, est objectivement fort différent de celui de l'écrit, à un point qui s'observe rarement dans les idiomes possédant une tradition d'écriture phonologique censée noter ce qui se prononce. Par là, on l'a vu, la diglossie française diffère de celle qui sépare l'arabe classique ou standard des arabes dialectaux, également bien distincts, mais telle que l'un et les autres se parlent, chacun avec sa grammaire propre.

Elle diffère aussi, bien évidemment, des cas où la différence des langues en concurrence saute trop aux yeux pour qu'on puisse la dénier, comme il en allait du latin face au « vulgaire » pendant tout le Moyen Âge européen (voire plus tard). S'agissant du français, la différence se voit efficacement camouflée par l'orthographe et par l'opinion commune quoique peu explicitée selon laquelle celle-ci note la vérité de la langue, que le parler déforme. C'est que la différence socialement valorisée n'est pas celle qu'une oreille ou un œil objectifs seraient capables de percevoir, mais celle qu'impose un consensus implicite non remis en cause, le plus souvent déterminé par des rapports de domination. C'est un tel consensus qui fait que, là encore contre toute évidence, l'occitan ou le créole martiniquais sont encore vus par beaucoup, y compris par ceux et celles-là mêmes qui les parlent, comme des dialectes ou des patois, donc des variétés inférieures de *la* langue... française, cela va de soi.<sup>176</sup> C'est pourquoi j'ai employé à plusieurs reprises et emploierai encore le terme « incarnation » afin de souligner au possible cette différence de nature (au moins en matière de flexion) entre FP et FO. Je serai donc amené à répéter bien des choses déjà abordées dans les chapitres précédents. Il est toutefois à espérer que leurs rapprochements les éclaireront d'une lumière nouvelle. Commençons par les similitudes.

### ***13.1. Similitudes entre FP et FO en matière de flexion***

Quant à la flexion nominale, FP et FO partagent le trait de définitude : sauf cas particuliers, les mêmes dans les deux incarnations (proverbes, locutions figées, énumérations), tout SN doit être marqué comme défini ou indéfini au moyen d'un déterminant préposé à son premier élément. Les descriptions fondées surtout ou seulement sur le FO n'y voient pas en général un fait de flexion, mais plutôt de composition syntaxique. Mais c'est qu'elles adoptent une conception trop étroite de la flexion, comme expliqué au chapitre 5. Dans le cadre plus libéral ici mis en œuvre, le caractère clitique des déterminants définis et indéfinis du FP les intègre de plein droit aux procédés flexionnels. Cela vaut aussi des déterminants déictiques et possessifs. De fait, FP et FO se caractérisent tous deux par la force des restrictions qui pèsent sur l'occurrence des noms sans déterminant.<sup>177</sup>

Est encore commune au FP et au FO la bipartition des lexèmes nominaux entre deux genres grammaticaux, l'un toujours non-marqué traditionnellement dénommé « masculin », l'autre parfois marqué et dit « féminin ».<sup>178</sup> Avec toutefois une différence : en FO, « le *e* final... favorise le féminin » (Melis & Godard 2021), et l'histoire l'explique, puisque cet <*e*>

<sup>176</sup> Dans *La Coscienza di Zeno* de Svevo, le héros se demande s'il doit faire sa demande en mariage « *in dialetto* » ou « *in lingua* ». La langue, c'est l'italien standard, le dialecte le frioulan, en réalité une autre langue. On en revient toujours à la définition d'Uriel Weinreich, écrivant à propos du yiddish : « Une langue est un dialecte avec une armée et une flotte » (« *A shprakh iz a dialekt mit an armey un a flot* »).

<sup>177</sup> Ces restrictions étaient moins fortes en moyen français et encore au dix-septième siècle (cf. Gougenheim 1951 : 63-64 ; Spillebout 1985 : 43).



correspond au /a/ final des féminins latins, occitans, espagnols, etc. Cet exposant graphique étant sans équivalent phonémique en FP (du moins dans sa variété « nordiste »), il s'ensuit que, malgré certaines corrélations statistiques entre tels phonèmes finaux et tel genre grammatical — p.ex. que les noms simples terminés par /ɛ̃/ et /ʒ/, tels *parpɛ̃* ‘parpaing’ et *menaz* ‘ménage’, les seconds montrant un <e> final à l'écrit, sont masculins à respectivement 99% et 94,2% (cf. Tucker, Lambert & Rigault 1977 : 68-125 ; cf. aussi Corbett 1991) — l'assignation d'un nom simple (non dérivé) à l'un ou l'autre genre n'est le plus souvent indiquée que par la forme du déterminant singulier associé, puisque les déterminants définis et indéfinis pluriels sont épïcènes. Voilà qui distingue le FP (et dans une moindre mesure le FO, ainsi que les autres langues d'oïl) du reste des langues romanes où l'on observe une corrélation assez régulière, quoique jamais absolue, entre la terminaison des noms simples et leur genre grammatical : ainsi en occitan languedocien où « *Los noms en -a son sentits coma femenins, mentre que los noms en -e, -i e de finala consonantica son sentits coma masculins* » (Alibèrt 2000 : 45). Le genre grammatical est par contre plus prédictible en FP comme en FO pour les noms construits pourvus de certains suffixes : tous les noms en *-sjɔ̃* / <-tion> sont féminins (cf. Melis & Godard 2021).

Je laisse de côté, car n'affectant pas le système, les discordes entre FP et FO en matière de genre grammatical des noms simples (cf. Gadet 1992 : 58-59). L'imprédictibilité de celui-ci les explique suffisamment. Qu'un nom simple commence par une voyelle suffit pour que l'article défini élidé n'en indique plus le genre grammatical, qui peut dès lors varier historiquement, voire en synchronie, sous l'effet de l'analogie ou sans raison décelable. Le problème ne date pas d'hier : au dix-septième siècle, *l=otɔn* ‘l'automne’ était féminin (Spillebout 1985 : 27) ; le mot est à présent masculin, sans doute aligné sur les autres saisons. À l'inverse, l'huître était masculine, elle est passée au féminin. *Aprèmidi* ‘après-midi’ a toujours les deux genres grammaticaux, au goût de chacun. Selon Gadet (1992 : 58) « en cas d'hésitation, un mot commençant par une voyelle est interprété comme féminin ». La remarque importante, me semble-t-il, est précisément que les cas d'hésitation sont peu nombreux. En général, le déterminant indéfini *ɛ̃/yn* ‘un.e’ décide sans ambiguïté. S'il y a quelques problèmes avec les noms à initiale vocalique, ce peut être dû à la relative débilite de la distinction phonétique entre /ɛ̃/ et /y/ dans, par exemple, /ɛ̃nylsɛʁ/ ‘un ulcère’ et /ynylsɛʁ/ ‘une ulcère’ (l'exemple est de Gadet, qui invoque ici une « prononciation archaïque »).

Nonobstant opacité et variations, la solidité de la catégorie du genre grammatical dans toutes les incarnations du français ne fait pas de doute. Le démontre entre autres choses le léger agacement que suscite toute erreur de la part d'un locuteur non-natif, anglophone par exemple, et le mal que l'on a à se retenir de le corriger. Gadet (1992 : 59) souligne à juste titre la force de la tendance à doter d'une forme féminine analogique, quand le besoin s'en fait sentir, les noms et adjectifs qui n'en possèdent a priori pas : cf. *rigolo* ‘rigolo’ / *rigolɔt* ‘rigolote’, *rikiki* ‘riquiri’ / *rikikit* ‘riquirite’. Il arrive aussi qu'une variation de genre grammatical se corrèle à une différence de sens : *legym* ‘légume’ n'est féminin que dans l'expression figée *grɔs legym* ‘grosse légume’, qu'on n'emploiera pas pour désigner un gros navet. Là, il s'agit du même lexème, pris au sens premier ou figuré. Mais on peut aussi avoir affaire à des lexèmes distincts tels *yn=mãf* ‘une manche’ et *ɛ̃=mãf* ‘un manche’, dont la distinction est aussi radicale que s'ils différaient phonologiquement. (On compte environ soixante-dix paires de ce type.)

Pour autant qu'il puisse se marquer, l'accord en genre grammatical des adjectifs attributs — a fortiori des épithètes — est aussi stable en FP qu'en FO : cf.

<sup>178</sup> J'emploie « non-marqué » en un sens strictement morphologique. L'idée, souvent agitée par les opposants à la « féminisation », selon laquelle le masculin serait sémantiquement neutre, donc bon pour les deux genres, se heurte aux travaux psycholinguistiques qui démontrent que la forme masculine des noms de référence humaine (p.ex. *chirurgien* vs. *chirurgienne*) introduit presque toujours un biais quant au sexe du référent.

*Se=z=yitr(ə) pɔrtygez etε trε=freʃ* ‘Ces huîtres portugaises étaient très fraîches’. Je doute qu’aucun.e francophone produise jamais *frε* au lieu de *freʃ* dans cet énoncé, sans parler de *pɔrtyge* ‘portugais’ au lieu de *pɔrtygez*. Si cela arrive, ce sera sous l’effet d’une confusion, toujours possible, corrigée dès que reconnue.

FP et FO s’accordent aussi quant aux critères sémantiques d’attribution de la pluralité aux noms comptables, c.-à-d. dénotant des entités présentes au monde sous forme d’individus discrets dénombrables. (Pour une étude approfondie de l’opposition comptable / massif, cf. Nicolas 2002, comme recommandé plus haut.) A un niveau qu’on peut qualifier de cognitif-pragmatique, tous deux appliquent le principe arithmétique expliqué au chapitre 6, avec pour conséquence qu’hormis certains contextes en FP (voir ci-dessous), tout nom dénotant un nombre de référents supérieur à 1 doit être pluralisé.<sup>179</sup>

S’agissant de l’usage des flexions verbales (non des formes mêmes), FP et FO ne diffèrent guère que sur le prétérit synthétique, quasiment inusité mais toujours compris dans le premier. Le passé du subjonctif est en déshérence dans l’un comme dans l’autre. Quant au présent du même mode, son emploi est sans doute plus variable en FP qu’en FO. Mais affirmer qu’il aurait disparu de celui-là est une grossière exagération. On entend parfois que le futur synthétique (*εl=fātəra* ‘elle chantera’) serait en voie de disparition, remplacé par le futur périphrastique (*εl=va fāte* ‘elle va chanter’). Un peu d’observation suffit à montrer que c’est tout à fait faux. En réalité, les deux formes ne sont ni sémantiquement ni pragmatiquement interchangeables : *wi, zə=ve fε-r la=vesel* ‘Oui, je vais faire la vaisselle’ et *wi, zə=fā-re la=vesel* ‘Oui, je ferai la vaisselle’ risquent fort de ne pas provoquer la même réaction de la part de l’interlocuteur/trice (cf. Co Vet & De Swart 2021).

### 13.2. Différences entre FP et FO en matière de flexion

En fait, l’écart entre FP et FO ne se situe pas au niveau des propriétés obligatoirement exprimées par les parties du discours soumises à flexion : définitude, genre grammatical et nombre pour les nominaux ; genre grammatical, nombre et cas pour les pronoms (intrinsèquement définis) ; TAM et accord PNG pour les verbes. Il se situe au niveau de l’expression de ces propriétés. On s’en souvient, on peut distinguer quatre modes combinables d’exposance selon le procédé morphologique mis en œuvre : synthétique (affixal ou non-linéaire, y compris le supplétisme), analytique (périphrases), semi-analytique (clitiques), défaut (expression par absence de marque). D’une façon générale, le FO recourt davantage que le FP à l’exposance synthétique affixale.

En FP définitude et genre grammatical se cumulent au singulier sur le mode supplétif : *lə* ‘le’ vs. *la* ‘la’, *ε* ‘un’ vs. *yn* ‘une’. De même en FO pour <le> et <la>. D’aucuns ont voulu décomposer ces formes en un morphème /l/ exprimant la définitude positive et un morphème vocalique /ə/ ou /a/ associé au genre grammatical. Voilà qui paraît hautement improbable, si du moins l’on se soucie de l’adéquation des analyses à la probable compétence des sujets parlants.<sup>180</sup> Le FO <une>, en revanche, ressortit plutôt au mode synthétique affixal pour autant que le <e> final, sans équivalent en FP, y constitue bien la marque du féminin. Autrement dit, et en se fondant sur la nature graphique du FO comme expliqué plus haut, on peut structurer

<sup>179</sup> Paradoxalement du point de vue arithmétique, 0 peut entraîner la pluralité : cf. *zero zurnal* ~ *zurno* ‘zéro journal ~ journaux’.

<sup>180</sup> On relève, il est vrai, une certaine corrélation entre la finale <a> ↔ /a/ et le féminin. Mais elle ne vaut que des noms propres, en particulier les prénoms féminins d’origine étrangère (Clara, Lydia, etc.) et certains noms géographiques — ce pour quoi, sans doute, le fleuve Volga, masculin au dix-neuvième siècle et sur les plaques de la « Rue du Volga » (Paris, 20<sup>e</sup> arrondissement), est passé au féminin (cf. « Les bateliers de la Volga »).

<une> comme ceci : <un<e>>. <sup>181</sup> Le FP *yn*, quant à lui, est sans structure interne et ne peut s'analyser que comme /yn/ supplétivement associé à *ē*. <sup>182</sup>

Au pluriel l'article défini est épïcène en FP comme en FO. Mais le mode d'expression diffère là encore. Le *le* 'les' du FP est supplétif par rapport à *lə* et *la*, sans plus de structure interne que *yn* ; le <les> du FO est formé synthétiquement par suffixation de <s>, marque de pluriel, à la base masculine : <le<s>>. A une étape antérieure du FP, quand /s/ faisait encore fonction de suffixe, l'article défini pluriel combinait apophonie et synthèse — et même double apophonie, vu qu'au début du dix-septième siècle on prononçait /le/ devant consonne et /ləz/ devant voyelle (Brunot 1939 : 195). Le nivellement de ces deux contextes et l'amuissement généralisé de /s/ final flexionnel ont aboli la composante synthétique, le suffixe remplacé par un exposant proclitique (voir chapitre 6), et transformé l'apophonie en supplétisme. L'identité du FP et du FO quant à l'article indéfini pluriel est par contre restée égale à elle-même : supplétisme pour FP *de* comme pour FO <des>, qui n'est pas le pluriel de <de> partitif, mais bien de <un<e>> (cf. Kupferman 1979 ; Tasmowski & Laca 2021).

L'expression du genre grammatical et du nombre des noms et des adjectifs est l'un des points sur lesquels FP et FO divergent le plus manifestement. Quant au nombre, le FO reste, on l'a vu, une langue romane occidentale typique, semblable à l'occitan languedocien, au catalan, à l'espagnol, au galicien et au portugais : le pluriel de <chat> est <chat<s>>, comparable, par exemple, à *gat* (/gat/) vs. *gats* (/gat-s/) (occitan languedocien, catalan). Un sous-ensemble des noms à finale <al> au singulier présente un radical à finale <au> propre au pluriel, dont la marque est alors <x> : <sup>183</sup> cf. <animal> vs. <animau<x>>, comparable au portugais *animal* (/eni'maɫ/) vs. *animais* (/eni'maj-f/). Le pluriel étant synthétiquement marqué sur le nom, la variation concomitante des déterminants et autres modifieurs relève de l'accord : <<La> <nuit> <tou<s>> <le<s>> <chat<s>> <sont> <gris>>. Accord qui ne touche pas les adjectifs terminés par un <s> (ou <x>) inhérent au radical, vu le rejet, et cela dès l'ancien français quand les consonnes finales se prononçaient encore, des consonnes redoublées ou longues (\* /gris:/) et, par le fait, d'écritures telles que \* <griss> (<gris<s>>). <sup>184</sup>

Le FP, par contre, s'écarte résolument des langues romanes, voire indo-européennes. Dans leur grande majorité, les noms et adjectifs y sont invariables quant au nombre, celui-ci indiqué (mais pas toujours) par les déterminants proclitiques associés. Il ne saurait alors être question d'accord : ce sont ces déterminants (*le* 'les', *se* 'ces', *me* 'mes', etc.) qui pluralisent le nom sur le mode semi-analytique. Quant aux noms masculins en /al/ de la classe de *zurnal* 'journal', leur forme en /o/ n'est pas à strictement parler un pluriel, mais plutôt un radical supplétif réservé aux SN pluriels — qu'ils peuvent, il est vrai, constituer à eux seuls dans les contextes qui autorisent les noms nus, non accompagnés de déterminants : cf. *zurno*, *livr*, *tu(t) ave bryle* 'Journaux, livres, tout avait brûlé'. Autrement dit, *zurno* (de même *šavo* 'chevaux', *kano* 'canaux', etc.) est une sorte de *pluralia tantum*, intrinsèquement pluriel même en l'absence d'un exposant de pluralité, que le /o/ final n'est pas en soi, mais seulement en tant que la forme globale *zurno* s'oppose paradigmatiquement à la forme globale *zurnal*. <sup>185</sup> On note que, même dans ces contextes, le FO maintient le marquage du nombre. En FP, hormis le

<sup>181</sup> Je rappelle que les parenthèses angulaires encadrent des graphèmes, susceptibles de représenter des exposants graphiques.

<sup>182</sup> Ou *ǣ* chez les francophones qui ont /ǣ/ dans leur inventaire vocalique. Cela ne change rien au caractère supplétif de l'association.

<sup>183</sup> Comme on sait, ce <x> remonte à une abréviation des scribes médiévaux pour la suite de lettres <us>, quand le pluriel d'*animal* s'écrivait *animaws* et se prononçait /animaws/. Étendu à quelques mots en <ou> (<choux>, <hiboux>, etc.), sa plus récente apparition semble être au pluriel de <ripou>, <ripoux>, à en croire l'affiche d'un film de 1984 qui eut son heure de succès. L'orthographe est passée dans l'usage.

<sup>184</sup> En espagnol, le pluriel de *gris* (/gris/) est *grises* (/ 'grises/), un /e/ épenthétique séparant les deux /s/, solution que le français n'a jamais adoptée.

<sup>185</sup> Un cas analogue, sauf pour la totalité et la singularité du supplétisme, est la paire russe *čelovek* 'humain, personne' vs. *ljudi* 'personnes, gens'.

supplétisme /al/ vs. /o/, ces noms nus sont non-spécifiés quant au nombre. Ou bien le contexte aide à décider, comme dans le vers de La Fontaine *Vo, va/...* ‘Veaux, vaches...’ — encore que Perette eût bien pu ne rêver qu’à *un* veau, *une* vache, etc. Ou bien on reste dans le vague : parlé, le titre du roman de Dostoïevski *krım ε fatimā* ne nous dit pas si l’auteur a voulu traiter d’un crime et d’un châtement ou du crime et du châtement en général — comme le laisse entendre le titre écrit <Crime et châtement>, traduction d’un russe sans ambiguïté singulier —<sup>186</sup> ou bien de plusieurs crimes et châtements particuliers ou des crimes et des châtements en général. Il n’est sûrement pas indifférent que ces contextes soient stylistiquement marqués. Je ne suis pas sûr qu’un corpus de FP en offrirait beaucoup d’exemples.

Cela étant, le FP n’ignore pas l’accord en nombre avec un adjectif épithète au sein d’un SN pluriel. Mais à condition que l’adjectif soit du type de *zenjal* ‘génial.e’ / *zenjo* ‘généiaux’ et le nom-tête masculin : cf. *de=poem zenjo* ‘des poèmes géniaux’, accord ; *de=rezyлта banal* ‘des résultats banals’, pas d’accord ; *de=z=ide zenjal* ‘des idées géniales’, pas davantage. Et de même entre un SN sujet pluriel et un adjectif attribut : cf. *se=poem sō zenjo* ‘Ces poèmes sont géniaux’, *se=rezyлта sō banal* ‘Ces résultats sont banals’, *se=z=ide sō zenjal* ‘Ces idées sont géniales’.

Enfin, le FP s’est doté d’un procédé original, quoique non dénué d’équivalents dans d’autres langues, de marquage du nombre : le *z* migrateur, cet exposant proclitique attachable à tout successeur vocalique dans la portée d’un opérateur pluriel, éventuellement implicite comme dans l’exemple de Morin (2005 : 10) : (*š=prā kwa*) *kōm z=afēr ?* ‘(On prend quoi) comme z-affaires ?’. Si l’opérateur est un déterminant, l’attache est (quasi) obligatoire et, sous l’influence de l’orthographe, traditionnellement décrite comme une liaison : *le=z=ami* ‘les amis’ — description historiquement exacte, puisque *z* tire son origine du suffixe *-s* de l’ancien et du moyen français, que le FO a conservé. Dans tout autre cas où *z* apparaît hors des contextes de liaison possible, son insertion est aléatoire et stigmatisée, d’autant moins probable que plus voyante : elle saute aux yeux dans *sēk=z=āfā* ‘cinq z-enfants’, elle se fait plus discrète dans *de=siber=z=atak* ‘des cyber-z-attaques’. L’opprobre n’entache pas la réalité du phénomène, il explique sa relative rareté.

S’agissant du marquage du genre grammatical, on a vu plus haut que le français en général, y compris le FO, s’est beaucoup éloigné des autres langues romanes. Le FO n’en a pas moins conservé de la romanité un suffixe <e> réalisé /ə/ du temps qu’il se prononçait — en gros jusqu’au milieu du dix-septième siècle — hérité du suffixe latin *-a* : cf. *cattus* ‘chat’ <chat> vs. *catta* ‘chatte’ <chatt<e>>. Plus intéressant pour notre objet est le fait qu’en FO, les adjectifs terminés en <e> sont invariables en genre : cf. *un résultat stupide, une idée stupide*. La raison en est assez limpide. Dans les adjectifs, le graphème <e> est à l’évidence la marque du féminin : cf. *vert* <vert> vs. *verte* <vert<e>>, *banal* <banal> vs. *banale* <banal<e>>. L’adjoindre aux items comportant déjà un <e> final au masculin entraînerait une suite de deux <e> que l’orthographe moderne répudie : *\*stupidee* <stupide<e>> (alors qu’elle tolère très bien <ée>, voire <éée> dans *créée* <cré<é<e>>>). En ancien et moyen français, quand les schwas finaux étaient prononcés et l’usage des accents fluctuant, l’obstacle était phonologique, à savoir l’impossibilité des suites de deux schwas : *\*/stypidə-ə/*.

La chute générale des schwas finaux en FP moderne a conduit à une situation différente. C’est désormais la quasi totalité de adjectifs qui ne varie pas en genre : *stypid* ‘stupide’, *banal* ‘banal.e’, *grēk* ‘grec.que’, *mējær* ‘meilleur.e’, etc. Reste un sous-ensemble variable, mais dont la flexion repose non plus sur la suffixation, mais sur un supplétisme partiel mettant en jeu deux formes qu’on peut qualifier de courte et longue respectivement : *āsjē* ‘ancien’ vs. *āsjen* ‘ancienne’, *blā* ‘blanc’ vs. *blāf* ‘blanche’, *bō* ‘bon’ vs. *bōn* ‘bonne’, *lō* ‘long’ vs. *lōg* ‘longue’, *pāti* ‘petit’ vs. *pātit* ‘petite’, *ru* ‘roux’ vs. *rus* ‘rousse’, etc. La consonne allongeante — parfois deux comme dans *syspē* ‘suspect’ vs. *syspēkt* ‘suspecte’ —

<sup>186</sup> *Prestuplenie i nakazanie*. Le pluriel serait *Prestuplenija i nakazanija*.

est le plus souvent imprévisible (voir chapitre 7). Il en va de même dans les noms simples dénotant des animés dont la différence sexuelle est exprimée par une paire court-long : *lu* ‘loup’ vs. *luv* ‘louve’, *rənar* ‘renard’ vs. *rənard* ‘renarde’, *fa* ‘chat’ vs. *fat* ‘chatte’, etc. Dire cette consonne « latente » est, me semble-t-il, une façon d’infiltrer la flexion du FO dans le FP.

Le FP apparaît ainsi plus entropique que le FO. Il ne faudrait toutefois pas exagérer ce qui risque d’être vu comme une déféctuosité. D’une part, même quand la forme longue ne se laisse pas prédire depuis la courte, l’une et l’autre restent unies par la constance de leurs patrons associés :  $[...]C^+V\#] \leftrightarrow [...]C^+VC\#]$ . Autrement dit, forme courte et forme longue se terminent par une même syllabe ou consistent en cette même syllabe (comme indiqué par le soulignement), si ce n’est que la syllabe de la première est ouverte, tandis que celle de la seconde est fermée. Il y a là une régularité qui devrait aider l’acquisition de ces paires, finalement plus transparentes que les paires entièrement supplétives telles que *buk* ‘bouc’ vs. *fəvr* ‘chèvre’, *frer* ‘frère’ vs. *sær* ‘sœur’, *fəval* ‘cheval’ vs. *zymā* ‘jument’, etc. D’autre part, comme montré au chapitre 7, les cas sont quand même assez nombreux en types et fréquents en occurrences où la consonne fermante est prévisible : p.ex. *āsjē* ‘ancien’ vs. *āsjen* ‘ancienne’ et tous les noms et adjectifs à finale /jē/ au masculin.

La principale différence entre FP et FO quant à la flexion pronominale est en fait corrélative de leur différence en matière de flexion verbale. Dans la mesure où les flexions synthétiques du FO, quoique moins riches que celles des autres langues romanes, distinguent encore assez bien personnes et nombres, les pronoms personnels sujets y ont le statut de mots autonomes, syntaxiquement équivalents à des SN et souvent sémantiquement redondants : cf. *tu chantes* <<tu> <chante<s>>>, où le trait [PNG 2SG] est marqué deux fois, par le pronom et par le suffixe <s>. Que les pronoms sujets soient obligatoires en français, et cela depuis fort longtemps, démontre simplement l’absence de corrélation nécessaire entre richesse de la flexion et « sujet nul » (voir chapitre 9). En FP, le rétrécissement de la flexion synthétique de personne-nombre suite aux bouleversements phonétiques du moyen français — alors que la flexion TAM s’est plutôt bien conservée — a fait de pronoms sujets déjà exigés des clitiques personnels : cf. *ty=fāt* ‘tu chantes’, qui ne se distingue que par l’indice de  $ʒə|ɛ||ʒ=fāt$  ‘je|elle(s)|on chante(nt)’. Ici, l’écart typologique entre FP et FO est à son plus large.

En revanche, les phénomènes les plus souvent remarqués comme séparant les deux systèmes en matière de flexion verbale, l’obsolescence du prétérit synthétique et la quasi extinction du passé du subjonctif, ne sont en rien propres au FP. En FO aussi, le second se fait rare et les domaines d’emploi du premier sont nettement circonscrits, limités en gros à la narration fictionnelle — évolution qui s’étend du reste au-delà du français, puisque la substitution du prétérit analytique au synthétique touche aussi l’espagnol européen, l’italien septentrional, le néerlandais, le haut-allemand et le yiddish.

Enfin, le FP ne montre, semble-t-il, aucune tendance à « simplifier » sa flexion TAM, du moins pour les verbes fréquents de la classe IIa qui abrite les « irréguliers ». Ceux-ci, en dépit des « fautes » relevées par les grammairiens, dont Frei (1929/2007) a bien démontré l’intérêt scientifique, préservent intactes pour l’essentiel les nombreuses alternances de bases, source majeure d’opacité. Aucun registre du FP, que je sache, ne conjugue \* $ʒə=dɔrm$  ou \* $nou=bwa$  à l’instar du créole haïtien qui dit *m(wen) domi* ‘je dors’ et *nou bwè* ‘nous buvons’. Par là, il est bien évident que FP et FO ne sont des langues distinctes que dans une certaine mesure, dans tels compartiments de leurs grammaires, tandis que français et créole haïtien le sont absolument.

On peut toutefois considérer que les alternances de bases remplissent des fonctions différentes en FP et FO dans certains contextes. Comparez  $Lə=fə dɔr$  ‘Le chat dort’ et  $Le=fə dɔrm$  ‘Les chats dorment’. En FO, les thèmes <<dor>> et <<dorm>> sont corrélés aux désinences <t> et <ent> qui expriment les traits [PNG 3SG] et [PNG 3PL] respectivement :

<dor<t>>, <dorm<ent>>. En FP, ces mêmes thèmes, /dɔr/ et /dɔrm/, expriment à eux seuls l'opposition de nombre à la 3<sup>e</sup> personne, flexion entièrement supplétive, donc, plutôt que combinant supplétisme et synthèse. Aux autres personnes, le FO montre cette même combinaison : <<je|tu> <dor<s>>>, <<nous> <dorm<ons>>>, <<vous> <dorm<ez>>>, tandis que le FP combine supplétisme, semi-analyticité et synthèse aux personnes 4 et 5 : *nu=dɔrm-ɔ̃*, *vu=dɔrm-e* ; supplétisme et semi-analyticité aux personnes 1 et 2 : *ʒə|ty=dɔr*, ainsi qu'aux personnes 3 et 6 quand le SN sujet est repris par un clitique personnel : *Lə=fa i(l)=dɔr* 'Le chat il dort', *Le=fa i(l)=dɔrm* 'Les chats ils dorment'.

### 13.3. Conclusion

Au total, la caractéristique du FP qui le sépare le plus nettement du FO (et des autres langues romanes) est l'importance du supplétisme, c.-à-d., je le rappelle, l'expression des distinctions flexionnelles par des formes à la fois différentes et inanalysables. Aucun département de la flexion qui ne le manifeste. En flexion nominale, le marquage du nombre par les déterminants est supplétif : *lə / la* vs. *le*, *ɛ̃ / yn* vs. *de*, *mɔ̃ / ma* vs. *me*, etc. — sauf à vouloir attribuer une signification à /l/ et /m/, démarche esthétiquement séduisante, mais, à mon avis, cognitivement irréaliste. (J'y reviens au chapitre suivant.) En FO, il ne l'est vraiment que pour l'indéfini, dans la mesure où /de/ diffère entièrement des formes du singulier — mais n'exprime pas à lui seul la pluralité, puisque <des> se laisse analyser <de<s>>, <de> radical supplétif, <s> suffixe de pluriel. De même <mes> analysé <me<s>>, <me> radical supplétif. Quant à <les>, on a vu qu'on peut y voir une forme synthétiquement fléchie <le<s>>, sans supplétisme graphique. Supplétive aussi en FP la forme de pluriel en /o/ des nominaux masculins en /al/ appartenant à cette classe, p.ex. *ɛ̃=zurnal / de=zurno*, tandis qu'en FO <journaux> admet l'analyse <journau<x>>, où <journau> est un radical second associé à la désinence <x>.

La disparition de /ə/ comme marque de féminin fait que la flexion en genre grammatical du FP est entièrement supplétive pour les noms et adjectifs biformes : cf. *lu / luv*, *ver / vert*, *āsjɛ̃ / āsjen*, la consonne additionnelle elle-même plus ou moins prévisible. En FO, en revanche, les deux derniers exemples sont synthétiques : <vert> / <vert<e>>, <ancien> / <ancienn<e>>. <sup>187</sup> Seul le premier met en jeu une base seconde associée à <e>, <louv> au lieu de <loup>.

Enfin, la préservation des désinences en FO réduit la part du supplétisme qui ne touche en gros que les quelques verbes qui le manifestent dans toutes les langues romanes, voire indo-européennes : <être>, <avoir>, <aller>. Pour le reste, on a plutôt affaire à des bases alternant au sein d'un espace thématique structuré par les valeurs de TAM, comme expliqué au chapitre 9. On a vu cette distinction illustrée au paragraphe précédent : supplétisme dans *dɔr* vs. *dɔrm*, bases alternantes dans <dor<t>> vs. <dorm<ent>> ; de même *bwa* vs. *bwav* et <boi<t>> vs. <boiv<ent>>. Les deux phénomènes coexistent en FP : cf. *el=par* 'elle part' et *el=part* 'elles partent', supplétisme ; *nu=part-ɔ̃* 'nous partons', base alternante. Autrement dit, il y a supplétisme quand la forme inanalysable exprime à elle seule une valeur flexionnelle, en l'occurrence le singulier du présent de l'indicatif pour *par* et la personne 6 du même présent de l'indicatif pour *part*. Cette même forme /part/ fait fonction de base alternante /part-/ si la valeur s'exprime par ailleurs comme dans *part-ɔ̃*.

Le caractère généralement non-spécifié pour le nombre des noms sans détermination (*vo*, *vaf*...) et, quant au nombre et au genre grammatical, de la plupart des adjectifs est une autre caractéristique du FP que ni le FO ni les autres langues romanes ne partagent. (Seul

<sup>187</sup> En moyen français, le doublement du <n> indiquait le maintien de la nasalité de la voyelle : on prononçait /āsjɛ̃n/. De nos jours, ce n'est plus qu'un procédé orthographique sans réelle justification. On pourrait très bien écrire <anciène>. On peut donc parler de « graphome » par analogie avec les « morphomes », morphes sans contenu sémantique (cf. chapitre 5, note 52).

l'occitan provençal s'en rapproche.) On se gardera toutefois de confondre cette non-spécification avec l'absence de la catégorie *grammaticale* de nombre que l'on observe dans des langues comme le chinois mandarin et le japonais — dont les locutrices et locuteurs (mais faut-il vraiment le préciser ?) sont parfaitement capables de distinguer un seul de plusieurs, c'est seulement que la grammaire de leurs langues, en ne leur imposant pas les procédés flexionnels pour ce faire, ne les contraint pas à exprimer une propriété dont il faut convenir qu'elle est souvent ou bien sans intérêt ou bien évidente dans la situation. En cas de besoin, le lexique contient les quantificateurs nécessaires (équivalents de « beaucoup », « quelques », etc.).

En revanche, nous (on permettra à l'auteur de s'inclure) ressentons cette contrainte. (Cette phrase le démontre, il est encore possible à l'écrit de séparer le clitique personnel du verbe. Avec « nous » et « vous », formes clitique et forte identiques, cela me semble également faisable en FP, moyennant une prosodie adaptée ; plutôt moins, en revanche, avec  $z\partial=$  et  $ty=$ . Mais il faudrait vérifier expérimentalement et/ou sur corpus, en particulier pour  $i(l)=$  et  $el=$ .) Les indices qui semblent prouver l'efficacité de cette contrainte ont été donnés. Je les rappelle. Premièrement, hormis les cas particuliers, les noms sont toujours marqués pour le nombre en FP, car accompagnés de déterminants fléchis pour la propriété, et cela que l'expression en soit ou non nécessaire ou simplement utile du point de vue informationnel. En d'autres termes, on dit  $z=em\_le=fa$  'J'aime les chats', parce que la composante flexionnelle de la grammaire du français, FP comme FO, ne permet pas de dire \*  $z=em\_fa$  'J'aime chat(s)', qui serait l'expression sans doute la plus adéquate, la moins ambiguë — et peut-être la plus répandue dans les langues du monde — du fait que mon affection se porte génériquement à l'espèce féline chantée par Baudelaire,<sup>188</sup> et non nécessairement à tel ou telle de ses représentants (dont aucun ne partage ma vie à l'heure où j'écris).<sup>189</sup>

Deuxièmement, le  $z$  migrateur décrit au chapitre 6 constitue à l'évidence un procédé pour indiquer *explicitement* une pluralité décodable, mais qui, vu les ressources flexionnelles du FP, ne s'exprime jamais par un exposant dédié semblable au /s/ audible de l'occitan languedocien, de l'espagnol, etc. : dans  $de=siber=z=atak$  'des cyber z-attaques', la forme du déterminant indéfini pluriel est supplétive. Dans  $kat(r\partial)=z=\tilde{a}f\tilde{a}$  'quatre z-enfants', la pluralité est avant tout sémantico-arithmétique. Il y a bien sûr le voisinage attractif de  $d\partial(z)$  'deux' et  $trwa(z)$  'trois'. Mais s'agit-il seulement d'une copie phonologique, ou bien plutôt d'une assimilation du /z/ « latent » de  $d\partial(z)$  et  $trwa(z)$ , originellement sans signification flexionnelle,<sup>190</sup> au  $z$  migrateur de  $gr\tilde{a}=z=\tilde{a}f\tilde{a}$  'grands enfants', lui-même vestige du /s/ pluriel audible de l'ancien français ?

Enfin, il arrive que l'équipement flexionnel se révèle incapable de faire apparaître le nombre — cf.  $o=p\partial pl$  'au peuple' ou 'aux peuples' — et cela suscite souvent un sentiment de gêne, de manque, auquel on remédie par des commentaires métalinguistiques (« au pluriel », « avec un s »). Il serait intéressant de savoir, si cela se peut, dans quelle mesure l'emploi de ces remèdes dépend de l'éducation scolaire. Surtout, la gêne éprouvée en dépend-elle ? Si elle est le symptôme d'une discordance entre besoin et moyen d'expression, le besoin naît-il de la compétence linguistique en FP ou bien de l'enseignement du FO ? Bref, il semble se

<sup>188</sup> « Non, il n'est pas d'archet qui morde / Sur mon cœur, parfait instrument, / Et fasse plus royalement / Chanter sa plus vibrante corde, / Que ta voix, chat mystérieux, / Chat séraphique, chat étrange, / En qui tout est, comme en un ange, / Aussi subtil qu'harmonieux ! » (*Le chat*).

<sup>189</sup> La moins ambiguë, parce que  $z=em\_le=fa$  peut aussi, selon le contexte, faire référence à des chats particuliers, et que  $z=em\_l\partial=fa$  'J'aime le chat' est triplement ambigu : générique (rarement), spécifique et (qui vient peut-être le plus facilement à l'esprit) culinaire.

<sup>190</sup> Ce n'est pas tout à fait exact. En ancien français, 2 et 3 se déclinent : le cas sujet masculin est *dui* /dʷi/ ou *doi* /doj/ ~ /dwe/, le cas régime masculin et le féminin sujet et régime *dous* /dus/ ou *deus* /dɔs/ ; de même *trei* ou *troi*, *treis* ou *trois*. Mais les formes non fléchies /dɔ(z)/ et /trwe(z)/ sont acquises dès le moyen français. A partir de 4, les numéraux ont toujours été invariables.

confirmer qu'en matière de flexion nominale, FP et FO diffèrent au niveau de l'exposance, semi-analytique ou synthétique, plutôt que par les catégories flexionnelles dont la grammaire impose l'expression. Il serait bon, néanmoins, d'étudier davantage les points soulevés dans les deux paragraphes précédents.

J'ai à de nombreuses reprises dans cette étude fait référence au moyen français. C'est que les propriétés caractéristiques du FP, qui le séparent du FO, trouvent pour la plupart leur origine dans les bouleversements qui affectent la langue qui fut celle de Montaigne et Rabelais — ou plutôt qui, commencés sans doute bien plus tôt, ont alors atteint un état d'avancement tel qu'ils se montrent dans la littérature et suscitent les commentaires des observateurs — en particulier l'amuïssement des consonnes suffixales (/s/ de pluriel, /t/ de personne 3 à l'imparfait, etc.) et du schwa marque de féminin. Bouleversement est le mot juste, car ce phénomène, qui touche toutes les langues d'oïl et dont la cause ultime nous échappe, apparaît comme le principal responsable de la mutation typologique que va connaître l'incarnation orale du français et à laquelle l'incarnation écrite, protégée par une orthographe conservatrice, échappe, mais qu'elle laisse transparaître.

L'événement est évidemment moins tragique que bien d'autres qui lui sont plus ou moins contemporains (guerres de religion, expulsion des Juifs d'Espagne et du Portugal, conquête et asservissement du Nouveau Monde, début de la traite négrière). Il n'en a pas moins son importance, et les conséquences s'en font encore sentir, la chasse aux « fautes » d'orthographe en est une. On en tirera deux leçons.

Premièrement, on voit là l'illustration on ne peut plus claire du caractère néo-darwinien du changement linguistique : une mutation fortuite — pour autant qu'on en ignore les causes — et limitée entraîne des modifications au niveau de l'organisme entier. Avec toutefois une différence entre le vivant et le linguistique. Quant au premier, la pression de l'environnement tend à sélectionner les mutations génétiques pourvues d'une valeur adaptative, en sorte que les individus qui les portent, survivant et se reproduisant davantage, finissent par représenter l'espèce à eux seuls. A l'évidence, un processus analogue s'est produit dans la langue : la mutation a été transmise, jusqu'à ce que grandisse une génération dont les membres ne prononçaient plus que /lemyr/ ce qui continuait à s'écrire — mais par et pour peu de gens — <les murs>.

Mais pourquoi les « mutantes » et « mutants » auraient-elles/ils engendré des progénitures plus abondantes que celles et ceux qui s'obstinaient à dire /lemys/ ?<sup>191</sup> Le « rendement » du système, c'est-à-dire sa capacité à exprimer les contrastes ressentis comme nécessaires, n'est pas en cause. J'ai montré, je l'espère, que la flexion semi-analytique du FP réussit aussi bien que la synthétique du FO à exprimer l'opposition de nombre des noms et adjectifs, de même que les propriétés TAM et PNG des verbes, sauf dans quelques cas — *o* 'au' ou 'aux', *i(l)/el=parl* 'il/elle parle' ou 'ils/elles parlent'. Cela étant, la seule hypothèse qui vient à l'esprit, pour aventurée qu'elle soit, est l'action de la sélection sexuelle et culturelle, l'une partagée entre tous les vertébrés (pour le moins), l'autre propre à l'espèce Homo (semble-t-il), dont Darwin (1871/2013) a démontré la nécessaire contribution, aux côtés de la sélection naturelle, à l'évolution des espèces (cf. aussi Stringer 2012/2014).

Le processus — dont il serait inutile de cacher que bien des points nous échappent, sans doute pour toujours — a pu se dérouler ainsi :<sup>192</sup> une fois déclenché par une cause inconnue et présent chez un nombre significatif de sujets parlants, l'amuïssement a introduit une différenciation au sein de la communauté linguistique. C'est un trait, trop souvent nocif, des humains qu'il n'est pas, ou guère, de différences indifférentes, toutes sont évaluées en

---

<sup>191</sup> On sait que la chute du /s/ des déterminants précède de beaucoup celle du /s/ des noms.

<sup>192</sup> Je m'en tiens, pour simplifier, au /s/ de pluriel. Il est probable que des processus analogues ont dû affecter les autres consonnes finales amuïes, simultanément ou avec un décalage temporel. Je renvoie à Labov (2010) pour un exposé bien plus approfondi.



bien ou en mal. Il s'est trouvé — là encore on ne sait pourquoi — que la prononciation /lemyr/ s'est vue socialement valorisée. (Elle aurait pu, a priori, se heurter à une stigmatisation qui l'aurait éliminée ou empêchée de se répandre.) Dès lors elle a conféré à celles et ceux qui la pratiquaient un prestige qui s'est exercé de deux façons : par une extension de leurs réseaux sociaux — davantage d'amis et connaissances susceptibles d'être influencées — et par un attrait supérieur en tant que partenaires sexuels aptes à engendrer des enfants à qui transmettre la forme prestigieuse. (Car le rôle du langage dans le choix, si choix il y a, des partenaires sexuels n'est plus à démontrer — cf. les références ci-dessus... et sa propre expérience. Faut-il préciser que tout ce processus ne fut ni délibéré, ni probablement accessible à la conscience de ses protagonistes ?) Comparer un beau parleur à un paon faisant la roue est un cliché, mais plus profond qu'il n'y paraît. N'y aurait-il pas alors quelque chose d'ironique à ce qu'une « faute » d'orthographe, cause de stigmatisme, ait son origine dans ce qui fut une séduction ?

J'en viens à présent à la deuxième leçon. Pendant que tout cela s'accomplissait, l'orthographe n'était ni fixée, ni sacralisée, ni surtout diffusée dans une population francophone dont la majorité était illettrée. La discordance entre l'oral et l'écrit ne pouvait donc être source de conflit. Tant qu'il ne s'agissait que de consigner par écrit des paroles ou des pensées, les graphies pouvaient varier sans inconvénient et, surtout, sans réprimande dans certaines limites. Par conséquent, et bien que les normalisations de l'orthographe soient évidemment bien antérieures — on les constate dès l'ancien français — il est permis de considérer sans trop d'exagération la notion de « faute » comme un effet de la scolarisation universelle de la fin du dix-neuvième siècle dans un contexte d'imposition d'une langue unique et « républicaine », dotée d'une écriture intangible, à un pays encore multilingue et de constitution politique encore indécise.

Cette imposition, qu'on la déplore ou l'approuve, il importe de comprendre qu'elle a recouvert deux entreprises généralement confondues. Dans le roman d'Alain-Fournier, les élèves de M. Seurel parlent français entre eux dans la cour de récréation. Il est plus probable qu'ils se parlaient en solognot. Par là, je veux rappeler que, le plus souvent, les instituteurs et institutrices des années 1880 se trouvaient face à des élèves dont la langue maternelle n'était pas le français, mais une autre langue d'oïl (un « patois »), ou une langue d'oc (y inclus le catalan) ou une langue non gallo-romane (alsacien, basque, breton, corse, flamand, j'en oublie sans doute). Leur tâche consistait donc en effet à leur enseigner le français, FO et FP *en même temps* : d'une part, leur apprendre à « mettre des s » au pluriel des noms et des adjectifs ; d'autre part, du même mouvement, leur enseigner une langue où ces s ne se prononcent pas, tâche sans doute facile avec des élèves de langue d'oïl, mais plus ardue avec des occitanophones ou des bretonnants.

Voilà qui nous ramène à ce qui a été le fil conducteur de cet ouvrage, à savoir la situation paradoxale de deux objets linguistiques, le français orthographique et le français parlé, très clairement différents l'un de l'autre quant à leurs composantes flexionnelles, et en même temps indissociables, en cela que la flexion du second est la seule réalisation possible de la flexion du premier, laquelle n'a d'existence que graphique. Pourrait-on résoudre ce paradoxe ? Faut-il souhaiter le faire ? Je n'ai pas de réponse à ces questions ni aucune légitimité pour vouloir en proposer une. Mais je pense qu'il est nécessaire d'être conscient du paradoxe, de l'accepter en tant que tel et, par là, de cesser de considérer le FO comme le « vrai » français, dont le FP ne serait que la production appauvrie.

## **Bibliographie**

Abeillé, Anne. 1993. *Les nouvelles syntaxes. Grammaires d'unification et analyse du français*. Paris : Armand Colin.

- Abeillé, Anne & Danièle Godard. 2021. *La Grande grammaire du français*. Arles : Actes Sud.
- Ackerman, Farrell, Gregory T. Stump & Gert Webelhuth. 2011. Lexicalism, periphrasis, and implicative morphology. In Robert D. Borsley & Kersti Börjars (eds), *Non-Transformational Syntax. Formal and Explicit Models of Grammar*, 325-358. Oxford : Wiley-Blackwell.
- Alibèrt, Loïs. 2000. *Gramatica occitana segon los parlars lengadocians*. Barcelona-Tolosa : Institut d'Estudis Occitans / Institut d'Estudis Catalans.
- Anonyme. 1997. *Bescherelle. La conjugaison pour tous*. Paris : Hatier.
- Arantéle. [www.arantele.org](http://www.arantele.org).
- Atkinson, Quentin D., Andrew Meade, Chris Venditti, Simon J. Greenhill & Mark Pagel. 2008. Languages evolve in punctuational bursts, *Science* 319, 588.
- Barra-Jover, Mario. 2009. Comment évolue un trait grammatical : le pluriel en français dans une perspective romane, *Romance Philology* 63, 25-67. halshs-00266121.
- Barra-Jover, Mario. 2010. Le pluriel nominal en français : un parcours sans faute(\*s), *Langue française* 168, 35-52.
- Barwise, Jon & Robin Cooper. 1981. Generalized quantifiers and natural language, *Linguistics & Philosophy* 4, 159-219.
- Beaulieu, Charles. 1927. *Histoire de l'orthographe française*. Paris : Champion.
- Bec, Pierre. 1973. *Manuel pratique d'occitan moderne*. Paris : Picard.
- Bickels, Balthasar & Johanna Nichols. 2013a. Fusion of selected inflectional formatives. Dryer, M.S. & M. Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology (<http://wals.info/chapter/26>).
- Bickels, Balthasar & Johanna Nichols. 2013b. Exponence of selected inflectional formatives. Dryer, M.S. & M. Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (<http://wals.info/chapter/26>).
- Blanche-Benveniste, Claire. 1997. *Approches de la langue parlée en français*. Paris : Ophrys.
- Blanche-Benveniste, Claire. 2002. Structure et exploitation de la conjugaison des verbes en français contemporain, *Le français aujourd'hui* 139, 11-22.
- Blanche-Benveniste, Claire. 2004. Singulier et pluriel en français parlé contemporain, *Bulletin de la Société de Linguistique de Paris* 99, 129-154.
- Blanche-Benveniste, Claire & André Chervel. 1969/1978. *L'orthographe*. Paris : Maspéro.
- Blevins, James P. 2006. Word-based morphology, *Journal of Linguistics* 42, 531-573.
- Boissin, Henri. 1975. *Grammaire de l'albanais moderne*. Paris : Chez l'auteur.
- Bonami, Olivier. 2021a. La morphologie et le lexique. Abeillé & Godard, Introduction-3.2.
- Bonami, Olivier. 2021b. Modes, temps et personnes du verbe. In Abeillé & Godard, chap. II-1.3.
- Bonami, Olivier & Gilles Boyé. 2003. Supplétion et classes flexionnelles, *Langages* 152, 102-126.
- Bonami, Olivier & Gilles Boyé. 2005. Construire le paradigme d'un adjectif, *Recherches linguistiques de Vincennes* 34, 77-98.
- Bonami, Olivier & Gilles Boyé. 2013. La nature morphologique des allomorphies conditionnées : les formes de liaison des adjectifs en français. Dans Georgette Dal & Dany Amiot (réd.), *Sillexicales 3 : Les unités morphologiques*. <https://stl.univ-lille.fr/productions-scientifiques/revues-et-collections-de-stl/sillexicales/>.
- Bonami, Olivier & Gilles Boyé. 2019. [Paradigm uniformity and the French gender system](#). In Matthew Baerman, Oliver Bond & Andrew Hippisley (eds), *Perspectives on Morphology: Papers in honour of Greville G. Corbett*, 171-192. Edinburgh: Edinburgh University Press.

- Bonami, Olivier, Gilles Boyé & Jesse Tseng. 2005. Sur la grammaire des consonnes latentes, *Langages* 158, 89-100.
- Booij, Geert. 1996. Inherent versus contextual inflection and the Split Morphology Hypothesis, *Yearbook of Morphology 1995*, 1-16.
- Bourdieu, Pierre. 1983. Vous avez dit « populaire » ?, *Actes de la recherche en sciences sociales* 46, 98-105.
- Boyé, Gilles. 2000. Problèmes de morpho-phonologie verbale en français, en espagnol et en italien. Thèse de doctorat, Université Paris 7.
- Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford : Blackwell.
- Bresnan, Joan & Sam A. Mchombo. 1995. The lexical integrity principle : Evidence from Bantu, *Natural Language and Linguistic Theory* 13(2), 181-254.
- Brisset, Jean-Pierre.**
- Brunot, Ferdinand. 1939. *Histoire de la langue française des origines à 1900. Tome IV : la langue classique (1660-1715)*. Paris : Librairie Armand Colin.
- Bybee, Joan L. 1985. *Morphology : A study of the relation between meaning and form*. Amsterdam : John Benjamins.
- Bybee, Joan L. 1988. Morphology as lexical organization. M. Hammond & M. Noonan (eds), *Theoretical Morphology*, 119-141. New York : Academic Press.
- Bybee, Joan L. 2005. La liaison : effets de fréquence et de construction, *Langages* 158, 24-37.
- Bybee, Joan & Jean E. Newman. 1995. Are stem changes as natural as affixes ?, *Linguistics* 33, 633-654.
- Caron, Bernard. 2015. Hausa Grammatical Sketch. HAL-SHS 00647533v3.
- Caron, Gauthier. 2011. Les niveaux de langue en créole réunionnais. Mémoire de master, Université Paris-Diderot.
- Carstairs-McCarthy, Andrew. 2010. *The Evolution of Morphology*. Oxford : Oxford University Press.
- Cerquiglini, Bernard. 1991. *La naissance du français*. Paris: Presses Universitaires de France.
- Cerquiglini, Bernard. 2004. *La genèse de l'orthographe française (XIIIe-XVIIe siècles)*. Paris : Champion.
- Cerquiglini, Bernard. 2007. *Une langue orpheline*. Paris: Editions de Minuit.
- Chaudenson, Robert. 1974. *Le lexique du parler créole de la Réunion*. Paris : Honoré Champion.
- Chaudenson, Robert. 2003. *La créolisation : théories, applications, implications*. Paris: L'Harmattan.
- Chevillet, François. 1994. *Histoire de la langue anglaise*. Paris : Presses universitaires de France.
- Chevrot, Jean-Pierre, Céline Dugua & Michel Fayol. 2005. Liaison et formation des mots en français : un scénario développemental, *Langages* 152.
- Chiss, Jean-Louis & Jacques David. 1992. La règle orthographique : conceptions théoriques et stratégies d'apprentissage, *Langue française* 95, 6-26.
- Chomsky, Noam. 1957. *Syntactic Structures*. The Hague : Mouton & C°. (1969. *Structures syntaxiques* (trad. Michel Braudeau). Paris : Le Seuil)
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge (Mass.) : MIT Press. (1971. *Aspects de la théorie syntaxique* (trad. Jean-Claude Milner). Paris : Le Seuil).
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge (Mass.) : MIT Press.
- Co Vet & Henriëtte De Swart. 2021. L'interprétation des temps verbaux. In Abeillé & Godard, chap. XI-2.
- Cohen, Marcel. 1967. *Histoire d'une langue : le français (des lointaines origines à nos jours)*. Paris : Messidor / Editions sociales.
- Collart, Jean. 1967. *Histoire de la langue latine*. Paris : PUF.

- Collins, Chris & Richard Kayne. 2020. Towards a theory of morphology as syntax. Ms.
- Corbett, Greville G. 1991. *Gender*. Cambridge : Cambridge University Press.
- Corbett, Greville G. 2009. Canonical inflection classes. Montermini F.F., G. Boyé & J. Tseng (eds), *Selected Proceedings of the Sixth Décembrettes : Morphology in Bordeaux*, 1-11. Somerville (Mass.) : Cascadilla Proceedings Project.
- Côté, Marie-Hélène. 2005. La liaison : de la phonologie à la cognition, *Langages* 152, 66-78.
- Creissels, Denis. 2006a. *Syntaxe générale : une introduction typologique, Vol. 1*. Paris : Hermès Sciences – Lavoisier.
- Creissels, Denis. 2006b. *Syntaxe générale : une introduction typologique, Vol. 2*. Paris : Hermès Sciences – Lavoisier.
- Croft, William. 1991. *Syntactic Categories and Grammatical Relations*. Cambridge : Cambridge University Press.
- Crysmann, Berthold & Olivier Bonami. 2017**
- Culicover, Peter W. & Ray Jackendoff. 2005. *Simpler Syntax*. Oxford : Oxford University Press.
- Damourette, Jacques & Edouard Pichon. 1911-1927. *Des mots à la pensée. Essai de grammaire de la langue française*. Paris : D'Artray.
- Darwin, Charles. 1871. *The Descent of Man, and Selection in Relation to Sex*. London : Murray. [2013. *La filiation de l'homme et la sélection liée au sexe*. Paris : Champion Classiques]
- De Clercq, Karen. 2009. *What is Nanosyntax ?* Conférence à l'Université de Genève.
- Delattre, Pierre. 1966. *Studies in French and Comparative Phonetics*. La Haye : Mouton.
- Deleuze, Gilles. 1968. *Spinoza et le problème de l'expression*. Paris : Les Editions de Minuit.
- Dell, François. 1973. *Les règles et les sons. Introduction à la phonologie générative*. Paris : Hermann.
- Dell, François. 1983. L'acquisition des règles phonologiques facultatives, *Langue française* 58, 15-22.
- Descusses, Martine. 1986. *Le patois ardennais de Gespunsart*. Paris : SELAF.
- Desrochers, R. 1994. Les liaisons dangereuses : le statut équivoque des erreurs de liaisons, *Linguisticae investigationes* 18, 243-284.
- Dryer, Matthew S. 2013. Prefixing vs. Suffixing in Inflectional Morphology. Dryer, M.S. & M. Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Téléchargeable à <http://wals.info/chapter/26>)
- Dubois, Jean, Mathée Giacomo, Louis Guespin, Christiane Marcellesi, Jean-Baptiste Marcellesi & Jean-Pierre Mével. 1973. *Dictionnaire de linguistique*. Paris : Librairie Larousse.
- Durand, Jacques & Chantal Lyche. 2003. Le projet « Phonologie du Français Contemporain » (PFC) et sa méthodologie. E. Delais and J. Durand (éd.), *Corpus et variation en phonologie du français : méthodes et analyses*, 212-276. Toulouse : Presses Universitaires du Mirail.
- Eco, Umberto. 1994. *La recherche de la langue parfaite dans la culture européenne* [trad. Jean-Paul Manganaro]. Paris : Le Seuil.
- Ehrhart, Sabine & Melanie Revis. 2013. Tayo. Michaelis S.M., P. Maurer, M. Haspelmath & M. Huber (eds), *The Survey of Pidgin and Creole Languages, Vol. II*. Oxford: Oxford University Press.
- Encrevé, Pierre. 1988. *La liaison avec et sans enchaînement. Phonologie tridimensionnelle et usages du français*. Paris : Le Seuil.
- Fabra, Pompeu. 1941. *Grammaire catalane*. Paris : Les Belles Lettres.

- Fayol, Michel & Jean-Pierre Jaffré. 1992. L'orthographe : approches linguistiques et psycholinguistiques, *Langue française* 95, 3-5.
- Fillmore, Charles J. 1968. The Case for Case. In Emmon Bach & Robert T. Harms (eds), *Universals in Linguistic Theory*, 1-88. New York: Holt, Rinehart, and Winston.
- Fouché, Pierre. 1959. *Traité de prononciation française*. Paris : Klincksieck.
- Fouché, Pierre. 1967. *Morphologie historique du français, le verbe*. Paris : Klincksieck.
- Foulquié, Paul. 1962. *Dictionnaire de la langue philosophique*. Paris: Presses universitaires de France.
- Fradin, Bernard. 1993.**
- Fradin, Bernard. 2003. *Nouvelles approches en morphologie*. Paris: Presses universitaires de France.
- Fradin, Bernard & Françoise Kerleroux. 2003.**
- Fraser, Norman M. and Greville G. Corbett. 1997. Default in Arapesh. *Lingua* 103: 25-57.
- Frei, Henri. 1929/2007. *La grammaire des fautes*. Rennes : Ennoïa.
- Furet, François & Jacques Ozouf. 1977. *Lire et écrire : l'alphabétisation des Français de Calvin à Jules Ferry*. Paris : Les Editions de Minuit.
- Gadet, Françoise. 1992. *Le français populaire*. Paris : Presses Universitaires de France.
- Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan A. Sag. 1985. *Generalized Phrase Structure Grammar*. Oxford : Oxford University Press.
- Giusti, Emilio. 1981. *Contes créoles de Guinée Bissau*. Paris : Conseil international de la langue française (CILF).
- Glaude, Herby. 2012. *Aspects de la syntaxe de l'haïtien*. Paris : Anibwé.
- Glynn, Ian. 1999. *An Anatomy of Thought : The Origin and Machinery of the Mind*. London : Phoenix.
- Goffman, Erving. 1986. La condition de félicité, *Actes de la recherche en sciences sociales* 64, 63-78 ; 65, 87-98 [trad. Alain Kihm]
- Goody, Jack. 1977. *The Domestication of the Savage Mind*. Cambridge : Cambridge University Press. (1979. *La raison graphique. La domestication de la pensée sauvage*, trad. Jean Bazin et Alban Bensa. Paris : Editions de Minuit.)
- Gougenheim, Georges. 1938. *Le système grammatical de la langue française*. Paris : Bibliothèque du français moderne.
- Gougenheim, Georges. 1951. *Grammaire de la langue française du seizième siècle*. Lyon : IAC.
- Gould, Stephen Jay. 2006. *La structure de la théorie de l'évolution*. Paris : Gallimard.
- Govain, Renauld. 2017. *Le parler bolith : histoire et description*. Boston : JEBKA Editions.
- Grammont, Maurice. 1914/1963. *Traité pratique de prononciation française*. Paris : Librairie Delagrave.
- Green, Georgia M. 2011. Modeling grammar growth : Universal Grammar without innate principles or parameters. Robert D. Borsley & Kersti Börjars (eds), *Non-Transformational Syntax. Formal and Explicit Models of Grammar*, 378-403. Oxford : Wiley-Blackwell.
- Grevisse, Maurice. 2009. *Le petit Grevisse. Grammaire française*. Bruxelles : DeBoek – Duculot.
- Grevisse, Maurice. 1980. *Le bon usage. Avec des remarques sur la langue française d'aujourd'hui*. Paris-Gembloux : Duculot.
- Guzmán Naranjo, Matías & Laura Becker. 2017. Quantitative Methods in African Linguistics — Predicting Plurals in Hausa. <https://mguzmann89.gitlab.io/pdf/hausa-post.pdf>
- Gwilliams, Jean-Remi King, Alec Marantz & David Poeppel. 2022. Neural dynamics of phoneme sequences reveal position-invariant code for content and order, *Nature Communications* 13, 6606.

**Hagège**

- Hagemeyer, Tjerk, Philippe Maurer-Cecchini & Armando Zamora Segorbe. 2020. *A Grammar of Fa d'Ambô*. Berlin : De Gruyter Mouton.
- Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. K. Hale & S.J. Keyser (eds), *The View from Building 20. Essays in Linguistics in Honor of Sylvain Bromberger*, 111-176. Cambridge (Mass.): MIT Press.
- Harris, Zellig S. 1951. *Methods in Structural Linguistics*. Chicago : University of Chicago Press.
- Harris, Zellig S. 1976. *Notes du cours de syntaxe*. Paris : Le Seuil.
- Haspelmath, Martin & Andrea D. Sims. 2010. *Understanding Morphology (2<sup>nd</sup> edition)*. London : Routledge.
- Hockett, Charles F. 1954. Two models of linguistic description, *Word* 10, 210-234.
- Hofstadter, Douglas R. 1985. *Metamagical Themas : Questing for the Essence of Mind and Pattern*. London : Penguin Books.
- Humboldt, Wilhelm von. 1974. *Introduction à l'œuvre sur le kavi* (trad. Pierre Caussat). Paris : Le Seuil.
- Huot, Hélène. 2005. *La morphologie. Forme et sens des mots du français (deuxième édition revue et actualisée)*. Paris : Armand Colin.
- Karttunen, Lauri. 1977. Syntax and semantics of questions, *Linguistics and Philosophy* 1, 3-44.
- Kail, Michèle & Michel Fayol (dir.). 2000. *L'acquisition du langage, vol. 1 et 2*. Paris : Presses universitaires de France.
- Kaye, Jonathan D. & Jean Lowenstamm. 1984. De la syllababilité. F. Dell, D. Hirst & J.-R. Vergnaud (dir.), *Forme sonore du langage*, 123-155. Paris : Hermann.
- Keenan, Edward L. & Jonathan Stavi. 1986. A semantic characterization of natural language determiners, *Linguistics & Philosophy* 9, 253-326.
- Kihm, Alain. 1993. What is it that you said? A study of obligatory focalization in two Creoles and beyond. In Byrne F. & D. Winford (eds), *Focus and Grammatical Relations in Creole Languages*, 141-162. Amsterdam : Benjamins.
- Kihm, Alain. 1995. Tayo, the strange bird from New Caledonia : determiners and tense-aspect in Tayo and their implications for creolization theory, *Journal of Pidgin and Creole Languages* 10(2), 225-252.
- Kihm, Alain. 2005. Les langues créoles. Jean-Marie Hombert (dir.), *Aux origines des langues et du langage*, 390-427. Paris : Fayard.
- Kihm, Alain. 2018. Le syntagme nominal en kriol de Guinée-Bissau et en kinubi : comparaison morphosyntaxique et sémantico-pragmatique, *Faits de langue* 49(1)
- Kilani-Schoch, Marianne & Wolfgang U. Dressler. 2005. *Morphologie naturelle et flexion du verbe français*. Tübingen : Gunter Narr Verlag.
- Klausenburger, Jürgen. 1977. Deletion versus Epenthesis: intra-versus interparadigmatic arguments in linguistics, *Lingua* 42, 153-160.
- Klausenburger, Jürgen. 1978a. French Linking Phenomena: A Natural Generative Analysis, *Language* 54, 21-40.
- Klausenburger, Jürgen. 1978b. Liaison 1977: The Case for epenthesis, *Studies in French Linguistics* 1/2, 1-20.
- Koyré, Alexandre. 1973. *Etudes d'histoire de la pensée scientifique*. Paris : Gallimard.
- Kuhn, Thomas S. 1962. *The Structure of Scientific Revolutions*. Chicago : The University of Chicago Press. (1983. *La structure des révolutions scientifiques*, trad. Laure Meyer. Paris : Flammarion.)
- Kupferman, Lucien. 1979. L'article partitif existe-t-il ? *Le français moderne* 47(1), 1-16.
- Kupferman, Lucien. 1994. Du : un autre indéfini ? *Faits de langues* 4, 195-203.

- Kuryłowicz, Jerzy. 1964. *The Inflectional Categories of Indo-European*. Heidelberg : Carl Winter Universitätsverlag.
- Kuwae, Kunio. 1983. *Manuel de japonais, volume 1*. Paris : L'Asiathèque.
- Kwon, Song-Nim & Anne Zribi-Hertz. 2006. Pluriel et généricité nominale : les enseignements du coréen. C. Dobrovie-Sorin (sld.), *Noms nus et généricité*, 159-195. Saint-Denis : Presses universitaires de Vincennes.
- Labov, William. 2010. *Principles of Linguistic Change. Cognitive and Cultural Factors*. Oxford: Wiley-Blackwell.
- Labelle, Marie, & Lori Morris. 2011. The acquisition of a verbal paradigm: verb morphology in French L1 children. Ms. UQAM ([www.archipel.uqam.ca/.../Labelle-Morris](http://www.archipel.uqam.ca/.../Labelle-Morris))
- Lakoff, George. 1987. *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind*. Chicago : University of Chicago Press.
- Laks, Bernard. 1980. Différenciation linguistique et différenciation sociale : quelques problèmes de sociolinguistique française. Thèse de 3<sup>e</sup> cycle, Université de Paris VIII.
- Laks, Bernard. 1983. Langage et pratiques sociales : étude sociolinguistique d'un groupe d'adolescents, *Actes de la recherche en sciences sociales* 46, 73-97.
- Laks, Bernard. 2005. La liaison et l'illusion, *Langages* 158, 101-125.
- Langendoen, D. Terence & Paul M. Postal. 1984. *The Vastness of Natural Languages*. Oxford : Blackwell.
- Lazard, Gilbert. 2006. *Grammaire du persan contemporain*. Téhéran : Institut français de recherche en Iran.
- Lazard, Gilbert. 2015. Thèses pour la linguistique pure, *Bulletin de la Société de linguistique de Paris* CX(1), 445-459.
- Leclerc, Jean-Marc & Sèrgi Javaloyès. 2004. *Le gascon de poche*. Chennevières-sur-Marne : Assimil.
- Lévi-Strauss, Claude. 1962. *La pensée sauvage*. Paris : Le Seuil.
- Lufin, Xavier. 2013. Kinubi. Michaelis S.M., P. Maurer, M. Haspelmath & M. Huber (eds), *The Survey of Pidgin and Creole Languages, Vol. III*. Oxford: Oxford University Press.
- Mallet, Géraldine M. 2008. La liaison en français : descriptions et analyses dans le corpus PFC. Thèse de doctorat, Université Paris Ouest Nanterre La Défense.
- Mammeri, Mouloud. 1986. *Précis de grammaire berbère (kabyle)*. Paris : Editions Awal.
- Manfredi, Stefano & Sara Petrollino. 2013. Juba Arabic. Michaelis S.M., P. Maurer, M. Haspelmath & M. Huber (eds), *The Survey of Pidgin and Creole Languages, Vol. III*, 54-65. Oxford: Oxford University Press.
- Marouzeau, Jules. 1961. *Lexique de la terminologie linguistique*. Paris : Librairie orientaliste Paul Geuthner.
- Martinet, André. 1965. *Éléments de linguistique générale*. Paris : Armand Colin.
- Martinet, André. 1969. *Le français sans fard*. Paris : Presses universitaires de France.
- Martinet, André. 1986. *Des steppes aux océans. L'indo-européen et les « Indo-Européens »*. Paris : Payot.
- Martinet, Jeanne. 1989. Un traitement fonctionnel de la liaison en français, *Folia Linguistica* 22(3-4), 293-299.
- Massot, Benjamin. 2008. Français et diglossie: décrire la situation linguistique française contemporaine comme une diglossie: arguments morphosyntaxiques. Thèse de doctorat, Université Paris 8. [<http://tel.archives-ouvertes.fr/tel-00726999>]
- Matthews, Peter H. 1972. *Inflectional Morphology : A theoretical study based on aspects of Latin verb conjugation*. Cambridge : Cambridge University Press.
- Mel'čuk, Igor A. 1993. *Cours de morphologie générale. Première partie : le mot*. Montréal : Presses de l'Université de Montréal — CNRS Editions.

- Mel'čuk, Igor A. 1994. *Cours de morphologie générale. Deuxième partie : significations morphologiques*. Montréal : Presses de l'Université de Montréal — CNRS Editions.
- Mel'čuk, Igor A. 1996. *Cours de morphologie générale. Troisième partie : moyens morphologiques. Quatrième partie : syntactiques morphologiques*. Montréal : Presses de l'Université de Montréal — CNRS Editions.
- Melis, Ludo & Danièle Godard. 2021. La formation, le genre et le nombre des noms. In Abeillé & Godard, chap. IV-2.
- Michaud, Alexis & Bonny Sands. 2020. Tonogenesis. Mark Aronoff (ed.), *Oxford Research Encyclopedia of Linguistics*. Oxford : Oxford University Press.
- Miller, Philip. 1992. *Clitics and Constituents in Phrase Structure Grammar*. New York : Garland Publishing.
- Milner, Jean-Claude. 1979. Le système de la négation en français et l'opacité du sujet, *Langue française* 44, 80-105.
- Milner, Jean-Claude & François Regnault. 1987. *Dire le vers. Court traité à l'intention des acteurs et des amateurs d'alexandrins*. Paris : Editions du Seuil.
- Mirambel, André. 1939/2002. *Grammaire du grec moderne*. Paris : Klincksieck.
- Mithun, Marianne. 1999. *The Languages of Native North America*. Cambridge : Cambridge University Press.
- Moignet, Gérard. 1988. *Grammaire de l'ancien français*. Paris : Klincksieck.
- Molino, Jean. 1985. Où en est la morphologie ? *Langages* 78, 5-40.
- Morin, Yves-Charles. 1979. La Morphophonologie des pronoms clitiques en français populaire, *Cahier de linguistique* 9, 1-36.
- Morin, Yves-Charles. 1987a. Remarques sur l'organisation de la flexion des verbes français, *ITL Review of Applied Linguistics* 77-78, 13-91.
- Morin, Yves-Charles. 1987b. French data and phonological theory, *Linguistics* 25, 815-843.
- Morin, Yves Charles. 2005. La liaison relève-t-elle d'une tendance à éviter les hiatus ? Réflexions sur son évolution historique, *Langages* 158, 8-23.
- Morin, Yves Charles. 2008. Les yods fluctuants dans la morphologie du verbe français. B. Fradin (resp.), *La raison morphologique. Hommage à la mémoire de Danielle Corbin*, 133-154. Amsterdam : John Benjamins.
- Morin, Yves Charles & Jonathan D. Kaye. 1982. The syntactic bases for French liaison, *Journal of Linguistics* 18(2), 291-330.
- Morisset, René & Georges Thévenot. 1975. *Les lettres latines*. Paris : Editions Magnard.
- Nicolas, David. 2002. *La distinction entre noms massifs et noms comptables. Aspects linguistiques et conceptuels*. Louvain : Peeters.
- Nida, Eugene A. 1949. *Morphology : The Descriptive Analysis of Words*. Ann Arbor : University of Michigan Press.
- Nølke, Henning. 1996. Où placer l'adjectif épithète ? Focalisation et modularité, *Langue française* 111, 38-58.
- Nordlinger, Rachel & Joan Bresnan. 2011. Lexical-Functional Grammar : Interactions between morphology and syntax. In Robert D. Borsley & Kersti Börjars (eds), *Non-Transformational Syntax. Formal and Explicit Models of Grammar*, 112-140. Oxford : Wiley-Blackwell.
- Noyé, Dominique. 1974. *Cours de foulfouldé. Dialecte peul du Diamaré, Nord-Cameroun*. Maroua : Mission catholique / Paris : Librairie orientaliste Paul Geuthner.
- Nurse, Derek & Philippson, Gérard. 2006. *The Bantu Languages*. London: Routledge.
- Ogden, Charles K. & Ivor A. Richards. 1923. *The Meaning of Meaning*. London : P. Kegan.
- Ó Siadhail, Mícheál. 1980. *Learning Irish. An Introductory Self-Tutor*. Dublin : Dublin Institute for Advanced Studies.



- Palancar, Enrique L. & Jean-Léo Léonard. 2016. Tone and inflection : An introduction. E.L. Palancar & J.-L. Léonard (eds), *Tone and Inflection : New Facts and New Perspectives*, 1-12. Berlin : De Gruyter.
- Pană Dindelegan, Gabriela (coord.). 2010. *Gramatica de bază a limbii române*. București : Academia Română.
- Partchieva, Para & Françoise Guérin. 1997. *Parlons tchéchéne-ingouche. Langue et culture*. Paris : L'Harmattan.
- Partee, Barbara H. 2006. Richard Montague (1930-1971). K. Brown (ed.), *Encyclopedia of Language and Linguistics, Vol. 8*, 255-257. Oxford : Elsevier.
- Penny, Ralph. 2002. *A History of the Spanish Language*. Cambridge : Cambridge University Press.
- Perlmutter, David M (ed.). 1983. *Studies in Relational Grammar*. Chicago : Chicago University Press.
- Perrott, D.V. 1951. *Swahili*. London : The English Universities Press.
- Pétursson, Magnus & Emil H. Eyjólfsson. 1996. *Manuel d'islandais*. Paris : Klincksieck.
- Pinker, Steven. 1999. *L'instinct du langage* (trad. Marie-France Desjeux). Paris : Odile Jacob.
- Post, Brechtje. 2000. Pitch Accents, Liaison and the Phonological Phrase in French, *Probus* 12, 127-164.
- Postal, Paul M. 2004. *Skeptical Linguistic Essays*. Oxford : Oxford University Press.
- Ragon, Eloi. 1962. *Grammaire grecque*. Paris : J. de Gigord.
- Raynaud de Lage, Guy. 1967. *Manuel pratique d'ancien français*. Paris : Picard.
- Rey, Alain, Danièle Morvan & Gilles Firmin. 2013. *Le Robert de poche 2014*. Paris : Dictionnaires Le Robert.
- Riegel, Martin, Jean-Christophe Pellat & René Rioul. 1996. *Grammaire méthodique du français*. Paris : Presses universitaires de France.
- Rigault, André (dir.). 1971. *La grammaire du français parlé*. Paris : Hachette.
- Robins, Robert H. 1959. In defense of WP, *Transactions of the Philological Society* (Oxford University), 116-144.
- Rouveret, Alain. 2015. *Arguments minimalistes*. Lyon : ENS Editions.
- Rowlands, E.C. 1969. *Yoruba*. London : The English Universities Press.
- Rowlett, Paul. 2013. Do French speakers really have two grammars ?, *French Language Studies* 23, 37-57.
- Sabouraud, Olivier. 1995. *Le langage et ses maux*. Paris : Odile Jacob.
- Salvat, Josep. 1973. *Gramatica occitana*. Tolosa : Privat.
- Saussure, Ferdinand de. 1915/1982. *Cours de linguistique générale*. Paris : Payot.
- Sauvageot, Aurélien. 1949. *Esquisse de la langue finnoise*. Paris : Klincksieck.
- Sauvageot, Aurélien. 1982. *Premier livre de hongrois*. Paris : Publications orientalistes de France.
- Schane, Sanford A. 1967. La liaison et l'éliision en français, *Langages* 8, 37-59.
- Schane, Sanford A. & Georges Cohen-Boulakia. 1973. Sur le degré d'abstraction de la phonologie du français, *Langages*, 32, 27-38.
- Schuchardt, Hugo. 1882. « Kreolische Studien I : Ueber das Negerportugiesische von S. Thomé (Westafrika) », *Sitzungsberichte der philosophisch-historischen Classe der kaiserlichen Akademie der Wissenschaften* (100).
- Selkirk, Elizabeth. 1974. French liaison and the X' convention, *Linguistic Inquiry* 5, 573-590.
- Spencer, Andrew. 1998. Morphophonological operations. A. Spencer & A.M. Zwicky (eds), *The Handbook of Morphology*, 123-143. Oxford : Blackwell.
- Spillebout, Gabriel. 1985. *Grammaire de la langue française du XVIIe siècle*. Paris : Picard.
- Spinelli, Elsa & Fanny Meunier. 2005. Le traitement cognitif de la liaison dans la reconnaissance de la parole enchaînée, *Langages* 158, 79-88.

- Starke, Michal. 2009. Nanosyntax : A short primer to a new approach to language, *Nordlyd* 36(1), 1-6.
- Stringer, Chris. 2012. *Lone Survivors. How We Came to Be the Only Humans on Earth*. New York : Henry Holt and C°. [2014. *Survivants. Pourquoi nous sommes les seuls humains sur terre* (trad. A. Kihm). Paris : Gallimard.]
- Stump, Gregory T. 2001. *Inflectional Morphology : A theory of paradigm structure*. Cambridge : Cambridge University Press.
- Stump, Gregory T. 2019. The false dichotomy of abstractive and constructive morphology. Ms.
- Tasmowski, Liliane & Brenda Laca. 2021. La détermination et la quantification. In Abeillé & Godard, chap. V-1.
- Tesnière, Lucien. 1959. *Éléments de syntaxe structurale*. Paris : Klincksieck.
- Teyssier, Paul. 1976. *Manuel de langue portugaise (Portugal – Brésil)*. Paris : Klincksieck.
- The Christian Brothers. 1980. *New Irish Grammar*. Dublin : C.J. Fallon Ltd.
- Thompson, Laurence C. 1991. *A Vietnamese Reference Grammar*. Honolulu : University of Hawaii Press.
- Tranel, Bernard. 1976. A note of final consonant deletion in modern French, *Lingua* 39, 53-68.
- Tranel, Bernard. 1986. French liaison and extrasyllabicity. In O. Jaeggli & C. Silva-Corvalan (eds), *Studies in Romance linguistics*, 283-395. Dordrecht : Foris.
- Tranel, Bernard. 1990. On suppletion and French liaison, *Probus* 2, 169-208.
- Trépos, Pierre. 1956. Troisième partie : de la collection à l'unité, *Annales de Bretagne* 63(2), 217-278.
- Tucker G. Richard, Wallace E. Lambert & André Rigault. 1977. *The French Speaker's Skill with Grammatical Gender : An Example of Rule-Governed Behavior*. The Hague : Mouton.
- Väänänen, Veikko. 1963. *Introduction au latin vulgaire*. Paris : Klincksieck.
- Wray, Alison & George W. Grace. 2007. The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form, *Lingua* 117, 543-578.
- Xanthos, Aris<sup>1</sup>, Sabine Laaha<sup>2</sup>, Steven Gillis<sup>3</sup>, Ursula Stephany<sup>4</sup>, Ayhan Aksu-Koç<sup>5</sup>, Anastasia Christofidou<sup>6</sup>, Natalia Gagarina<sup>7</sup>, Gordana Hrzica<sup>8</sup>, F. Nihan Ketzrez<sup>9</sup>, Marianne Kilani-Schoch<sup>1</sup>, Katharina Korecky-Kröll<sup>2</sup>, Melita Kovačević<sup>8</sup>, Klaus Laalo<sup>10</sup>, Marijan Palmović<sup>8</sup>, Barbara Pfeiler<sup>11</sup>, Maria D. Voejkova<sup>12</sup>, & Wolfgang U. Dressler<sup>2</sup>. 2011. On the role of morphological richness in the early development of noun and verb inflection, *First Language* 31(4) : 461-479.
- Zemach, Eddy M. 1970. Four ontologies, *Journal of Philosophy* 67. [Repris dans F.J. Pelletier (ed.) 1979. *Mass Terms : Some Philosophical Problems*, 63-80. Dordrecht : Reidel.]
- Zink, Gaston. 1989. *Morphologie du français médiéval*. Paris : Presses Universitaires de France.
- Zink, Gaston. 1990. *Le moyen français*. Paris : Presses Universitaires de France.
- Zribi-Hertz, Anne. 2013. De la notion de grammaire *standard* dans une optique diglossique du français, *Journal of French Language Studies* 23(1), 59-85.
- Zwicky, Arnold M. 1977. *On Clitics*. Bloomington: Ohio State University.

