HAL
open science

# De la fécondité de certaines transgressions dans le domaine linguistique

Berthille Pallaud

## To cite this version:

Berthille Pallaud. De la fécondité de certaines transgressions dans le domaine linguistique . Voix Plurielles, 2015, Voix Plurielles, 21 (1), pp.167-185. hal-01153971

**HAL Id: hal-01153971**

**https://hal.science/hal-01153971v1**

Submitted on 20 May 2015

**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.

# De la fécondité de certaines transgressions dans le domaine linguistique

**Berthille PALLAUD,** Université-Aix-Marseille, UMR 7309 Laboratoire Parole et Langage, France

## Introduction

Il est impossible de prétendre rendre compte des effets des transgressions dans les domaines scientifiques tant il apparaît que la recherche de façon générale, ne peut avancer qu'en contestant des théories, des résultats et des faits antérieurement avérés. Il y eut, bien sûr, l'exemple fameux de l'astrophysique au dix-septième siècle avec Copernic, Galilée et bien d'autres physiciens depuis des siècles, qui montra l'importance d'une transgression majeure puisqu'elle remettait en cause un ordre dogmatique dans la conception de l'univers. Elle fut si essentielle qu'on la nomma plus tard « révolution copernicienne ». Dans le domaine linguistique, si certaines transgressions ont bien eu lieu, d'autres n'ont pu se manifester et bénéficier de leur fécondité. Ce propos sera traité à travers trois types d'exemples. Le premier est un point de vue sociolinguistique sur la politique linguistique française à l'égard des langues régionales et de la langue des signes : le cœur de l'argumentation étant le débat sur l'importance du bilinguisme. Dans le deuxième exemple, le français parlé est au centre des études sur la variation par rapport au français écrit. L'histoire de la constitution des grands corpus de français parlé témoigne des difficultés rencontrées pour rendre compte de cette variation. Le troisième exemple présente une étude de certaines caractéristiques « transgressives » du français oral.

## 1. Transgression et évolution des langues en France : la peur du bilinguisme

### *1. 1. Les émancipations n'ont pas toujours eu lieu*

L'histoire des langues en France nous apprend que l'une d'elles fut investie d'autorité par décision royale (dès le seizième siècle). Le français évoluant à partir du francilien (une variété des langues d'oïl) devint la langue de France et remplaça le latin dans tous les actes administratifs. C'est toujours le cas actuellement : selon Bernard Cerquiglini, en 1999, les langues de France sont au nombre de soixante-quinze[1]. Quant aux langues régionales de France, leur histoire est l'exemple d'une résistance linguistique qui capitula durant le vingtième siècle. Comme on le verra, cette capitulation a eu certains effets qu'on peut juger « régressifs » au regard du développement des communications entre les pays européens.

Selon Henriette Walter (16), la France a une particularité parmi les pays d'Europe : s'est perpétuée « une longue tradition qui lie la langue à la politique » (il suffit pour s'en convaincre de lire l'essai de Françoise Morvan paru en 2002). On retrouve là l'importance des institutions (politique, administrative, scolaire, judiciaire) qui, par leurs lois et leurs décrets, influencent les habitudes des populations et par là, en l'occurrence, le destin des langues. Cette situation explique la place particulière de la France, pays massivement monolingue, alors que l'Europe est majoritairement plurilingue (Bartolotti et Marian). Le tableau suivant (Walter 17) énumère ces décisions et illustre ce propos :

| **Langues régionales et Institutions** |
|---|
| 15 août 1539 : Ordonnance de Villers-Cotterêts : tous arrêts et procédures seront délivrés en *langage maternel et non autrement* (la langue française remplace le latin dans les écrits administratifs) |
| 4 juin 1794 : Rapport de l'abbé Grégoire sur *la nécessité et les moyens d'anéantir les patois et d'universaliser la langue française*. |
| Mai 1981 : 56ème des 110 propositions du programme de François Mitterrand : *les langues et cultures minoritaires seront respectées et enseignées.* |
| Avril 1999 : Mise en place d'un Observatoire linguistique émanant de la Délégation Générale de la Langue Française. |
| 23 juillet 2008 : Révision de la Constitution (art. 45-1) : *les langues régionales appartiennent au patrimoine de la France.* |

Henriette Walter rappelle qu'au début du vingtième siècle, les Français étaient au moins **bilingues** ; ils parlaient bien sûr leur parler régional, la langue du foyer, et utilisaient plus ou moins facilement d'autres parlers régionaux. Ils devaient également parler le français devenu langue administrative dès 1539 et langue de scolarisation depuis 1794. Cette année-là, le rapport de l'Abbé Grégoire (*Rapport sur la nécessité et les moyens d'anéantir les patois et la nécessité d'universaliser la langue française*) qui fut l'artisan, lors de la Révolution française, de cette stratégie à l'égard des langues de France, invoqua plusieurs arguments dont le principal était d'ordre politique. Au moment où la patrie républicaine se construisait, il était à craindre que les régions frontalières, profitant de la communauté linguistique avec les régions au-delà des frontières, ne cherchent à s'émanciper de la tutelle de la République et fassent sécession (Julia *et al*). Imposer le monolinguisme, en l'occurrence, l'usage exclusif du français pour tout acte écrit public, était la solution choisie pour pallier cette éventualité. Un siècle plus tard, à partir de 1880, avec les lois scolaires de Jules Ferry, l'école (obligatoire, gratuite et laïque) fut l'instrument de la lutte contre le bilinguisme oral qui caractérisait l'ensemble de la population française. Les langues régionales étaient toujours considérées comme source de danger. Le problème est que ces conceptions quasi religieuses de l'unicité linguistique liée à la nation et de la nocivité du bilinguisme durèrent en France deux siècles :

> L'Abbé Grégoire n'est pas seulement un régicide, il est aussi un abbé constitutionnel, il accepte la pluralité religieuse et il pose la langue comme facteur unifiant à la place du corps du roi et à la place de l'unicité religieuse. Va ainsi se développer en France une religion de la langue française. (Encrevé, P. 23)

L'école, ne reconnaissant que le français comme langue d'instruction orale aussi bien qu'écrite, imposa donc le français en interdisant strictement et de façon parfois humiliante l'emploi de la langue régionale (nommée alors « dialecte » voire « patois » en ces lieux). Par exemple, en Alsace, les personnes nées vers 1950 ont encore connu ces contraintes. La plupart des enfants naissaient dans un bain linguistique alsacien (langue germanique) et découvraient plus ou moins, vers 3 ou 4 ans, une deuxième langue, la langue française. En 1969, malgré cette situation de bilinguisme « additif », 91% des Alsaciens encore bilingues et se déplaçaient en Allemagne, Suisse et Autriche sans avoir besoin de l'anglais pour comprendre et se faire comprendre. Cette situation pouvait paraître enviable lors des colloques, à ceux qui, non alsaciens ou non germanophones, communiquaient, dans ces pays, par le truchement de l'anglais. Les récentes études (Phillip) révèlent qu'en 1998, seulement 50% des Alsaciens continuent de parler le dialecte alsacien. Or, non seulement, le bilinguisme précoce n'est pas nocif (Tabouret-Keller, « Nocivité », *Bilinguisme* ; Tabouret-Keller et Luckel) mais une chance (Hagège, 1996). Seul le foyer familial peut procurer à la petite enfance le privilège d'un plurilinguisme « profond », un bilinguisme précoce simultané ou consécutif. On pourrait donc dire que les Alsaciens sont en train de perdre un effet bénéfique qui, en partie, ne devait rien à l'éducation scolaire.

Un autre exemple, celui du Pays basque qui n'a pas eu la même histoire que l'Alsace dans son rattachement à la France, ce qui souligne le caractère exceptionnel de la résistance de la langue basque présente dans une partie de l'Aquitaine. Des écoles basques, ont été créées en 1969. La langue basque bénéficie en France d'enquêtes sociolinguistiques régulières depuis près de vingt ans. En 2011, sur les 31,6% de Basques natifs présents en Pays basque, 21,4% seulement sont bilingues actifs, 9,1% sont des bilingues passifs et 69,4% ne sont pas bascophones. On constate la même pente régressive observée en Alsace alors que comme l'écrit Henriette Walter, « 50% des gens interrogés dans cette région sont favorables à une promotion de la langue basque » (53).

La seule exception à ce déclin général des langues régionales en France est celui de la langue corse qui, en 1995, était comprise par 81% des habitants de l'île et parlée par 64% (134) : seule région française donc, où un bilinguisme langue française-langue régionale semble encore actif.

Et Henriette Walter de conclure, après Claude Hagège (« L'enfant » et *Combat*), non sans employer un conditionnel souligné, que ce bilinguisme précoce *aurait été* sans doute la meilleure préparation « pour l'apprentissage d'une langue de grande diffusion » (235). Le caractère avantageux du plurilinguisme précoce (dès 8 mois) est d'ailleurs rappelé par des travaux récents réalisés au Canada (Werker *et al*., a et b ; Bartolotti et Marian) : leurs résultats montrent des effets cognitifs bénéfiques si on les compare à ceux d'enfants nés en milieu monolingue.

Il y aurait donc des transgressions qui n'ont pu avoir lieu au siècle passé en France. On s'émeut actuellement de la disparition de ces langues régionales et de ce bilinguisme *à portée de main* en voie de disparition. Il est vrai qu'on tente de remédier à la situation en favorisant l'apprentissage d'une autre langue que la langue française dès la deuxième année d'école primaire (BO, 2007 et 2012).

***1.2 La langue des signes : le choix entre monolinguisme ou bilinguisme ?***

Toujours dans le cercle scolaire français, une autre langue, la langue des signes (LSF), a connu les contraintes liées aux variations dans la politique éducative des établissements scolaires. D'abord, il a été longtemps admis que la surdité rendait problématique voire impossible l'éducation scolaire des enfants sourds dans un établissement classique où ils devaient s'adapter à l'enseignement prévu et donné pour des enfants entendants (c'est-à-dire un enseignement essentiellement oralisé). Heureusement, dès 1760, l'abbé Charles Michel de l'Épée avait créé un Institut National des Jeunes Sourds pour enseigner les enfants sourds par le moyen d'une langue gestuelle (la LSF) fondée sur ses observations de leurs communications spontanées. Son initiative, poursuivie par l'abbé Sicard, sera abandonnée en 1880 par décision autoritaire du congrès international de Milan qui proscrit l'usage de la langue des signes afin de garantir à l'enfant sourd une intégration sociale pleine et entière au monde qui l'entoure : « Le congrès, considérant que l'usage simultané de la parole et des signes mimiques a le désavantage de nuire à la parole, à la lecture sur les lèvres et à la précision des idées, déclare que la méthode orale **pure** doit être préférée » (cité dans Encrevé, F., 89).

En France, la langue des signes est alors bannie de l'enseignement. Cette décision est lourde de conséquences quand on sait que 95% des enfants sourds profonds naissent de parents entendants (Encrevé F.). Ces derniers ignorent souvent l'existence de la langue des signes et ne pourront plus, dans ces conditions, l'apprendre eux-mêmes. Outre que l'enfant ne peut plus acquérir à l'école cette LSF, on le décourage de la pratiquer puisqu'elle l'éloigne de cet idéal que constituerait la langue orale française.

Heureusement, dès 1900, des professeurs transgressent ces directives et commencent à tolérer graduellement l'usage de la langue des signes mais uniquement comme communication entre élèves. La transmission des connaissances, quant à elle, se fait, toujours oralement. C'est ainsi que, malgré l'interdiction de signer en classe, la LSF ne disparaît pas, les enfants sourds la transmettant de génération en génération, la plupart du temps pendant la récréation. « Dès 1900, certes, les élèves sourds continuent à la pratiquer dans les grandes institutions, mais entre eux et de manière confidentielle. Elle n'est plus la langue de l'instruction, de la transmission des connaissances et de l'épanouissement intellectuel » (Encrevé F., 97).

Concernant la LSF, si les transgressions existent alors mais seulement dans la sphère privée, elles restent de peu de poids devant les directives imposées à l'enseignement. Les enseignants spécialisés qui enseignaient dans les Instituts pour enfants sourds se devaient de faire appliquer ces directives, d'autant plus que la technique des implants cochléaires réactivait la possibilité d'atteindre l'idéal d'une intégration de l'enfant sourd dans la sphère orale. Il reste que de nombreux sourds profonds refusent cette technique de l'implantation cochléaire qui, selon eux, n'aboutirait qu'à faire de mauvais entendants. Ils craignent de se retrouver rejetés aussi bien de la communauté des entendants que de celle des sourds. La propension au monolinguisme conçue dans une stricte alternative**,** *ou la langue des signes ou la voix*, a fait place plus tard au bilinguisme, *et la langue des signes et la voix* (ce qui avait été réclamé très tôt et en vain par la communauté des sourds). Ce n'est que dans les années 1970 que l'usage de la langue des signes va se développer à nouveau dans le cadre scolaire. Cependant l'interdiction d'enseigner par la méthode gestuelle perdure officiellement jusqu'en 1991 soit la date de l'amendement législatif signé par Laurent Fabius grâce auquel les parents d'enfant sourd et leur enfant peuvent choisir un enseignement soit exclusivement oraliste, soit *bilingue* (le français oral et la langue des signes française). Ce n'est que par la loi du 11 février 2005 qu'en France a été mis officiellement fin à l'obligation de la méthode orale pour l'éducation des sourds.

Peut-on souscrire à la comparaison faite par Florence Encrevé (92) qui note qu'à l'instar de la langue des signes, une langue régionale peut survivre dans la sphère privée ? Les chiffres avancés par Henriette Walter attestent plutôt du contraire. On peut certes interdire à un enfant de parler sa langue de communication (langue gestuelle ou régionale) mais la mise en demeure de percevoir et prononcer une langue orale pour qui est sourd ne peut être assimilée à celle qui impose un usage et un enseignement dans une langue inconnue (ou conçue comme telle). Dans le cas de la langue L2 (la seconde langue imposée), le handicap à

franchir n'est pas physiologiquement enraciné (comme c'est le cas dans la surdité profonde) ; tout au plus, peut-il l'être psychologiquement (comme ce peut l'être parfois pour des enfants vivant dans un milieu bilingue). C'est pourquoi, il semble que la langue gestuelle a un caractère de nécessité pour la survie du jeune enfant que n'a pas la langue régionale. La LSF peut être la langue d'entrée dans la communication humaine pour l'enfant sourd au même titre qu'une langue orale : « À condition d'une imprégnation précoce [de la langue des signes], l'enfant forme ses premiers signes, puis ses premières phrases, au même âge qu'un enfant entendant » (Busquet, et Allaire).

Quel que soit le point de vue et quelle que soit la méthode, il est entendu que l'enfant doit pouvoir entrer en communication avec son entourage le plus tôt possible. Pour un enfant sourd profond, la communication visuelle est la seule possible du moins dès les premiers instants de vie. Tout comme les enfants entendants (Boysson-Bardies), l'enfant sourd s'approprie activement les éléments de communication que son entourage lui fournit et qu'il peut percevoir.

Les avancées actuelles dans la technique des implants cochléaires[2] (en particulier, les appareils numériques) conduisent à nouveau à réexaminer l'usage et l'enseignement de la langue des signes y compris dans la sphère privée. Ces progrès techniques semblent encourager un dépistage néonatal de la surdité puisqu'ils permettent une implantation dès la première année de vie chez l'enfant sourd et assurer ainsi une meilleure intelligibilité de la parole[3]. Par là, on peut espérer un apprentissage précoce plus complet d'une langue orale (c'est l'avis de la CNNE, 1994). Il reste que rien n'autorise à penser que l'apprentissage précoce d'une communication visuelle (lecture labiale, mimique, langue française parlée complétée LPC et langue des signes) en parallèle avec une langue orale empêcherait cet apprentissage bilingue. « Des classes bilingues sont ouvertes dans différentes villes. Cette possibilité récente d'une éducation bilingue, même si elle n'est pas géographiquement accessible à tous, va dans le sens d'une meilleure prise en compte des besoins des personnes sourdes » (Busquet et Allaire). Le bilinguisme concernant la langue des signes et le français oral n'est donc plus transgressif puisque, actuellement, il est même recommandé.

## 2. Du pur à l'impur. Le développement des corpus oraux enregistrés, transcrits et annotés en France

Les bases documentaires, qu'elles soient orales ou écrites, ont été indispensables pour certaines études en linguistique. L'inventaire des corpus de français oral réalisés depuis le début du vingtième siècle (Cappeau et Seijido ; Cordereix sur le site de la DGLFLF) permet

de constater qu'il a d'abord été question d'enregistrer une langue française « une et pure » même si elle était cernée de variations diverses (dialectales, patoisantes, locales ou individuelles). Ce n'est que, dans un deuxième temps, que les variations du français oral ont été enfin reconnues dignes d'être conservées et donc enregistrables.

***2.1 Techniques d'enregistrement et pureté de la langue : les corpus historiques***

Le recours à des corpus oraux a été soumis dès le début aux contraintes et avancées de la technique puisque, pour conserver la parole dans sa modalité orale, il faut l'enregistrer à l'aide d'un appareil[4]. L'apparition du phonographe au début du vingtième siècle a suscité de nouvelles recherches en linguistique. C'est en 1899, à Vienne en Autriche que fut créée la première institution au monde, dévolue à l'enregistrement et à la conservation d'archives sonores (Cordereix). En France, c'est à l'initiative de Ferdinand Brunot, avec l'aide de l'industriel Emile Pathé, que furent inaugurées, le 3 juin 1911, les Archives de la Parole (la première pierre d'un Institut de Phonétique voulu par l'Université de Paris). Les corpus de Brunot enregistrés en France, entre 1912 et 1914, portent aussi bien sur le français (300 enregistrements) que sur les dialectes régionaux. La perspective concernant le français est la suivante : la recherche de *la parole pure, au timbre juste, au rythme impeccable, à l'accent pur* soit la description d'une unité modèle du français, *Le Français,* face aux dialectes encore massivement parlés et, cela va sans dire, aux paroles françaises pleines d'impuretés, au timbre douteux, aux rythmes fautifs, à l'accent impur, etc. Ces archives entièrement numérisées sont conservées au Département de l'audiovisuel à la Bibliothèque Nationale de France et consultables sur Internet[5].

Cette perspective puriste est encore présente 40 ans plus tard (au début des années 50). C'est l'époque d'une nouvelle invention technique, le magnétophone. Paul Rivenc, qui fut l'un des linguistes sollicités (avec Gougenheim, Michea et Sauvageot) pour le projet de recueil de français parlé, raconte que son premier enregistrement fut réalisé dans une épicerie grâce à un magnétophone à bandes magnétiques importé des Etats-Unis. Le gouvernement souhaitait améliorer la diffusion de la langue française dans le monde et finança l'établissement du Corpus de français fondamental réalisé dans les années 1951-1955. Ce corpus donna lieu à une vaste enquête destinée à l'enseignement du français aux étrangers et aux populations de l'Union française (Gougenheim *et al*.). Dans une perspective similaire, radio RFI produit tous les jours un Journal en Français Facile qui dure 10 min (sur la base de 300 mots) tout comme *La Voix de L'Amérique* diffuse des programmes depuis 1959 en *Special English* (1500 mots). De plus, cette enquête a permis d'établir une liste de mots et d'indications grammaticales[6]. Ce projet fut critiqué par ceux qui y voyaient une volonté

d'établir un sous-français pour classes inférieures et dénonçaient les relents de colonialisme déjà soulignés dans le projet anglais de *Basic English* pour les colonies anglaises. On considérait que ces enregistrements fournissaient un français abâtardi, indigne des efforts d'enseignement de la langue française. Comme le rappelle Rivenc (5), la tentative de description d'une grammaire de la parole, la première au monde, se heurtait aux linguistes et enseignants qui ne pouvaient accepter comme représentant la langue française des énoncés remplis de phénomènes caractéristiques de l'oral (effacement du *ne*, marqueurs de discours, énallages énonciatifs, etc). D'autres critiques furent formulées qui tenaient, en particulier, à la qualité médiocre des enregistrements. Traumatisé par les violentes polémiques suscitées par ce projet de recueil, Gabriel Gougenheim n'a pas voulu que les enregistrements réalisés en majorité sur disques de papier magnétisé et donc difficilement audibles sur le plan phonétique n'alimentent encore plus les critiques (Rivenc 6) : elles furent détruites à la demande de Gougenheim. Il ne nous reste que leurs transcriptions textuelles. Quoiqu'il en soit, le titre même de ce recueil « Corpus de français fondamental » énonce bien le souci d'établir et de conserver la trace d'un français de base ou mieux de référence. Pour preuve, dans un entretien en 2006, Rivenc, qui participa au projet du corpus de français fondamental, déclara : « Nous avions convenu de limiter nos enquêtes à des régions où le français n'était pas en contact avec des « langues régionales…. les modèles de référence restaient bien, secrètement mais fermement, ceux de la langue écrite » (4-5).

Par la suite, dix ans plus tard, les recueils de paroles ont bravé ce cadre imposé du bien parler français et se sont ouverts à la variation par deux types de recherche, l'une dans une perspective sociolinguistique à Orléans (le Corpus d'Orléans), l'autre dans une approche syntaxique à Aix-en-Provence (Corpaix et Corpus de Référence de Français Parlé). La brèche était ouverte pour de multiples études et projets de recueil de français parlé.

### *2.2 L'ouverture à la variation dans les corpus oraux. La pêche au harpon et la pêche au chalut*

Dix ans plus tard, donc, un autre corpus de français oral est constitué de 1968 à 1971 sur la base d'enregistrements de locuteurs vivant à Orléans (ESLO-1). Des enquêtes sociolinguistiques sont conduites par des universitaires britanniques (University of Essex) et en collaboration avec le BELC (Bureau pour l'Etude de l'enseignement de la Langue et de la Civilisation française de Paris). La visée de ces enquêtes est didactique (entre autres pour l'enseignement du Français langue étrangère). Trois cents heures d'interview ont été réalisées, soit en face à face soit dans d'autres situations (conversations téléphoniques, réunions publiques, transactions commerciales, repas de famille, entretiens médico-pédagogiques).

Cette perspective sociolinguistique (Bergougnioux) sera reprise, à partir de 2008, par une deuxième enquête à l'origine d'un second corpus (ESLO-2) qui est toujours ouvert actuellement.

A Aix-en-Provence, à partir de 1975, bien avant que l'avancée des techniques du numérique ne puisse le justifier et le faciliter, une nouvelle initiative a été entreprise sous l'impulsion de Jean Stéfanini et poursuivie par Claire Blanche-Benveniste. L'intérêt n'était plus la constitution d'un corpus pour la sauvegarde d'une langue pure et unique. Il n'y aurait pas un français oral mais des variétés orales de français. La question de la valeur d'une variété, de sa capacité d'autorité est suspendue car il est question de décrire des emplois, des usages. Le projet de cette équipe, le Groupe Aixois de Recherches en Syntaxe (G.A.R.S.), affirme que : « La variabilité chez un même locuteur est une donnée fondamentale qui ne peut se mesurer que dans la diversité des situations de parole » (Bilger et Blanche-Benveniste).

Ce projet comprenait également une approche psycholinguistique ; par exemple, la demande d'explication et la demande de narration font apparaître des phénomènes syntaxiques différents. Il tenait, donc, à multiplier les contextes de prises de paroles et les types de locuteurs. Ce corpus connu sous le nom de Corpaix dont certains extraits ont été publiés (Blanche-Benveniste, Rouget et Sabio) est composé de cassettes audio avec leur transcription dactylographiée, plus ou moins complète, sur support papier. On conçoit que, dans ces conditions, l'exploitation morphosyntaxique de ces transcriptions ait demandé un temps considérable ne serait-ce que pour relire et ensuite rassembler les faits et les éléments recherchés. La numérisation des enregistrements et des transcriptions n'existait pas encore, pas plus que les outils d'extraction de faits linguistiques qui ont été construits depuis. En 1995, cette équipe fut chargée d'explorer la variable géographique pour la langue française en métropole. Le recueil d'entretiens dans plusieurs régions de France (projet sous la responsabilité de Blanche-Benveniste) a bénéficié des avancées technologiques et a permis l'établissement du *Corpus de Référence de Français Parlé* remis à la DGLFLF (*Délégation générale à la langue française* et *aux langues de France*) par l'équipe DELIC qui succéda au GARS.

A partir de l'année 1993, les nouvelles technologies n'ont fait que permettre l'élargissement et l'amplification de la perspective du GARS et de l'approche sociolinguistique tenue dans le Corpus d'Orléans. Benzitoun, en 2011, dans son état des lieux sur les corpus de français parlé, signale les travaux des années 2000 qui revendiquent à leur tour : « La nécessité de prendre en compte des corpus diversifiés pour une meilleure approche

des faits langagiers dans leur ensemble et appellent à la constitution de ressources communes et partageables. »

Il reste que, comme le soulignent Taravella Eshkol *et al*, malgré ces axes de recherche pionniers, une volonté institutionnelle a fait défaut en France, ce qui explique le retard pris dans la constitution d'un gros corpus national comportant de multiples sources de variétés langagières. Comme, on le sait, la situation a été différente en Grande Bretagne (le BNC, *British National Corpus*), en Russie (*Russian National Corpus*) ou encore, aux USA (*Santa Barbara Corpus of spoken american English*). Si, par volonté institutionnelle, le français écrit a été le premier constitué en un grand corpus, le corpus *Frantext*, le français oral a du mal à sortir de la marginalisation (Cappeau et Gadet). Il reste que les nouvelles technologies, en rendant possible « la pêche au chalut » ont encouragé voire contraint, dans les dix dernières années, le développement de corpus de français oral. On a osé abandonner le terrain de la langue idéale écrite.

L'archivage de toutes sortes de corpus oraux de français parlé s'est répandu et les études de ces variétés ne cessent de se développer dans tous les niveaux de langue. La numérisation des données orales et vidéo, alignées sur leur transcription et les annotations de leur description fournissent une énorme quantité de données, ce qu'on appelle des corpus enrichis[7]. Une volonté institutionnelle est maintenant clairement manifestée dans les appels d'offres. A en juger par l'exemple du projet Ortolang sur ces ressources et dont l'objectif est d'étendre leur diffusion, leur valorisation et leur pérennité :

> La version 0 du portail Ortolang est maintenant ouverte (www.ortolang.fr). L'architecture matérielle de l'équipement est aujourd'hui opérationnelle et son architecture logicielle poursuivra son évolution jusqu'en en 2016, année où est programmée la phase de plein fonctionnement. « À terme, l'objectif est aussi de faire d'ORTOLANG, un nœud du réseau CLARIN (Common Language Resources and Technologies : www.clarin.eu) » conclut Jean---Marie Pierrel.

## 3. Les disfluences dans le langage oral : des transgressions vis à vis de l'ordre écrit

L'élaboration de l'oral par le locuteur a des caractéristiques qui sont souvent résumées par le terme de *transgression* : en particulier, transgression des bonnes règles du français correct c'est-à-dire du français tel qu'on l'écrit et non tel qu'on le parle. Ce français oral présente des inconvénients évidents pour les études TAL (Traitement Automatique des Langues). Ces écarts à la norme de l'écrit sont-ils des fantaisies individuelles ?

Rivenc raconte une anecdote à propos d'Aurélien Sauvageot qui, on l'a vu, fut à l'origine du projet de *Français Fondamental* :

> « Permettez-moi de citer une anecdote particulièrement éclairante : en prenant connaissance des premiers enregistrements recueillis, Sauvageot [pourtant à l'origine du projet d'enregistrement] déclara péremptoirement et jusque dans sa famille que les gens cultivés s'exprimaient tout autrement et beaucoup mieux, même dans leurs conversations les plus familières. C'est alors que son fils Serge, qui devait avoir une quinzaine d'années, et deviendrait plus tard l'excellent africaniste que l'on sait, s'empara du magnétophone paternel et le cacha sous la table au cours d'un repas animé qui réunissait autour de son père plusieurs linguistes renommés, et enregistra les conversations. Aurélien Sauvageot, beau joueur mais très perplexe, nous apporta cet enregistrement à Saint- Cloud, et chacun dut faire son mea-culpa et réviser ses hypothèses de linguiste. » (3)

En effet, le français oral a été longtemps considéré, surtout par les grammairiens, comme porteur de scories, de ratés et d'erreurs peu dignes d'intérêt. La grande fréquence de ces phénomènes, nommés maintenant *disfluences*, est attestée par de nombreuses études (Rossi et Defare ; Shriberg) et nécessite une reconnaissance précise. Mais qu'est-ce qu'une disfluence ? Ce mot emprunté à l'anglais s'est imposé dans les études qui sont consacrées à ces phénomènes. Définir ce qu'est une simple disfluence dans le langage oral d'un locuteur « standard » permet de préciser les caractéristiques d'un oral dysfluent (et non simplement disfluent) alors révélateur d'une pathologie (Barkat-Defradas *et al.* ; Starkweather ; Duez). Par exemple, la répétition spontanée et involontaire de syllabes ou de mots peut être un indice de dysfonctionnement, comme la *palilalie*[8]. Or la répétition est très fréquente dans les énoncés de nos corpus sans, pour autant, être signe de pathologies (Bigi *et al.*).

Par ailleurs, un intérêt nouveau pour ces disfluences tient à une demande des nouvelles technologies dans le domaine des énoncés oraux : pouvoir obtenir une transcription propre d'un énoncé oral, c'est-à-dire nettoyée de ses scories orales, afin que la morpho-syntaxe orthodoxe de l'énoncé soit rétablie et que puisse être décrypté le message informatif, par exemple, par des logiciels pour des entreprises très diverses (banques, enquêtes téléphoniques, répondeurs électroniques de tous ordres). De ce point de vue, l'énoncé oral transgresse les règles respectées le plus souvent, dans les énoncés écrits.

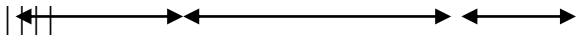
***L'axe de recherche : Approche morpho-syntaxique des disfluences***

Lors de conversations, les énoncés oraux se caractérisent par de nombreuses variations dans le rythme de la fluence verbale pouvant aller jusqu'à l'auto-interruption de mots ou de syntagmes, en passant par des perturbations morphosyntaxiques. Ces auto-interruptions et auto-variations dans la fluence verbale sont signalées, dans la grande majorité des cas, (mais pas dans tous), par l'apparition d'une ou de plusieurs sortes d'évènements: pauses silencieuses ou remplies (« euh »), items insérés au milieu d'un syntagme ou même d'un mot.

AB_620 C'était *euh tu vois* complètement loufoque **comme si-** *ouais euh* **comme situation**

Elizabeth Shriberg (1994) a décrit la structure sous-jacente à ces perturbations : le *Reparandum*, lieu de la production source rompue au point de rupture, suivi de l'*Interregnum* (moment potentiellement utilisé par le locuteur) et enfin du *Reparans* (lui aussi potentiel) qui correspond à la reprise du déroulement syntagmatique. L'avantage de cette description est qu'elle permet de distinguer ce qui précède l'interruption, ce qui lui succède éventuellement sans renouer le fil du discours et ce qui vient réparer, poursuivre ou laisser inachevée la construction syntaxique interrompue. L'intérêt de la catégorie *Interregnum*, sauf quand elle est vide, est qu'elle est composée d'éléments dont la liste n'est pas infinie. Leur présence permet de signaler et reconnaître une interruption dans le flux verbal quelle qu'en soit l'issue.

|*Reparandum* | *Interregnum* |*Reparans*|



tu perds | **un peu** |comment dire euh + | **un peu** |des repères

Le *Reparandum* est défini par les éléments que l'on retrouve dans le *Reparans*. Le *Reparans* (ce qui suit, en poursuivant, répétant, modifiant ou laissant inachevé) permet donc d'identifier le *Reparandum*. Le *Reparandum* comme le *Reparans* sont les deux éléments ou ensembles d'éléments qui encadrent l'Intereregnum. La rupture sur l'axe syntagmatique n'est pas toujours l'occasion d'une reprise plus ou moins partielle de l'énoncé même si cela varie selon les locuteurs. Deux sortes d'auto-interruptions sont distinguées et sont décrites pour en établir les différences et les similarités. Certaines de ces interruptions ne sont que **suspensives** (le moment de suspension passé, l'énoncé est simplement poursuivi), les autres sont dites **disfluentes** (le moment de suspension passé, l'énoncé n'est pas simplement poursuivi, il est « perturbé »). Ces perturbations sont soit des reprises partielles de ce qui a déjà été dit, soit des abandons laissant le texte elliptique :

Suspensives : AB_620 **c'était** *euh tu vois* **complètement** loufoque

YM _2432 tu le fais **avec** *euh* **ce** que tu veux

Disfluente avec reprise : YM_1506 un gosse **j'en** // **j'en** // **moi j'en** voulais pas

CM 884 *mais **les** euh **les** nanas du foyer elles étaient pas au courant*

Disfluente avec abandon : CM 1139 ah si mais **j'ai** // c'est un truc qui m'avait fait bien rire

Toutes les auto-interruptions et leurs effets ont été intégralement annotés dans les huit dilogues du corpus CID. En distinguant l'auto-interruption, ce qui la signale et les effets qu'elle provoque, cette étude a permis d'analyser les relations existant entre ces deux types d'événements : les interruptions simplement suspensives et celles, disfluentes, qui provoquent

une réparation ou un abandon et constituent des perturbations dans le déroulement de l'énoncé puisque, sur le seul point de vue syntaxique, l'élément qui succède à l'interruption n'est pas celui qui est attendu.

***Résultats partiels***

Les auto-interruptions suspensives diffèrent des auto-interruptions disfluentes sur plusieurs paramètres. Elles sont plus nombreuses. En moyenne, un peu plus de la moitié des interruptions observées (58%) ne sont que suspensives, les autres étant disfluentes et se subdivisant en deux sous-catégories : les reprises (29%) sont deux fois plus nombreuses que les abandons d'énoncés (13%). Que les disfluences avec reprise d'une partie de l'énoncé soient quantitativement supérieures aux abandons d'énoncé est un résultat qui a été observé également lors d'entretiens semi-directifs enregistrés (Pallaud et Henry ; Dister). Ces résultats obtenus sur 8 dilogues[9] du CID corroborent ceux qui avaient été trouvés dans une précédente étude pilote sur deux dilogues du CID (Pallaud et Rauzy) ; on peut les considérer comme robustes et représentatifs de la situation de dialogue conversationnel spontané.

***Le nombre moyen de mots sans interruptions***

Qu'elles soient suspensives ou disfluentes, les interruptions varient fortement en fonction des dilogues et des locuteurs. Leur fréquence représente en moyenne une rupture tous les 7,1 mots (écart-type : 1,246). Plus un énoncé est interrompu, plus le nombre de mots entre deux interruptions est faible. Lors de ces dialogues, le maximum de mots prononcés sans interruption est en moyenne de 10 mots, le minimum de 4 mots.

Les interruptions qui ne sont que suspensives sont les plus nombreuses : elles se produisent avec une fréquence moyenne de une tous les 12 mots (de 8 à 15 mots). Leur fréquence est une composante essentielle du débit de parole. Comme elles sont toujours suivies d'au moins une pause de 200 ms ou d'un marqueur de discours, plus elles sont nombreuses, plus le débit de parole diminue, puisque le débit tient compte du temps de parole total (pauses comprises).

Il n'en va pas même en ce qui concerne la fréquence des interruptions disfluentes. Elles sont en moyenne nettement moins nombreuses que les suspensives (une toutes les 19 mots) et varient beaucoup plus selon les locuteurs : de 8 à 33 mots (écart-type : 7,7). Même si les interruptions disfluentes contribuent également à diminuer le débit de parole, cette catégorie d'interruption diffère de la catégorie des interruptions suspensives : elle est moins fréquente et surtout nettement moins homogène.

Les interruptions disfluentes sont donc, plus que les interruptions suspensives, représentatives des caractéristiques énonciatives des locuteurs. Leur diversité, au moins dans

la variabilité de leur fréquence, rend mieux compte que les ruptures sans disfluence des possibilités énonciatives de chaque locuteur. De plus, le locuteur qui suspend son énoncé, le reprend en partie ou le modifie, est aussi susceptible de le laisser inachevé. Chacune de ces disfluences, en terme de fréquence, est caractéristique du locuteur : en effet, ces deux types de disfluences sont corrélés de façon significative (Fig.1 - Coefficient de corrélation : 0,56 ; ddl=15 ; p<.02). Le locuteur qui a une grande fréquence de disfluence avec reprise est aussi celui qui laissera souvent son énoncé inachevé.

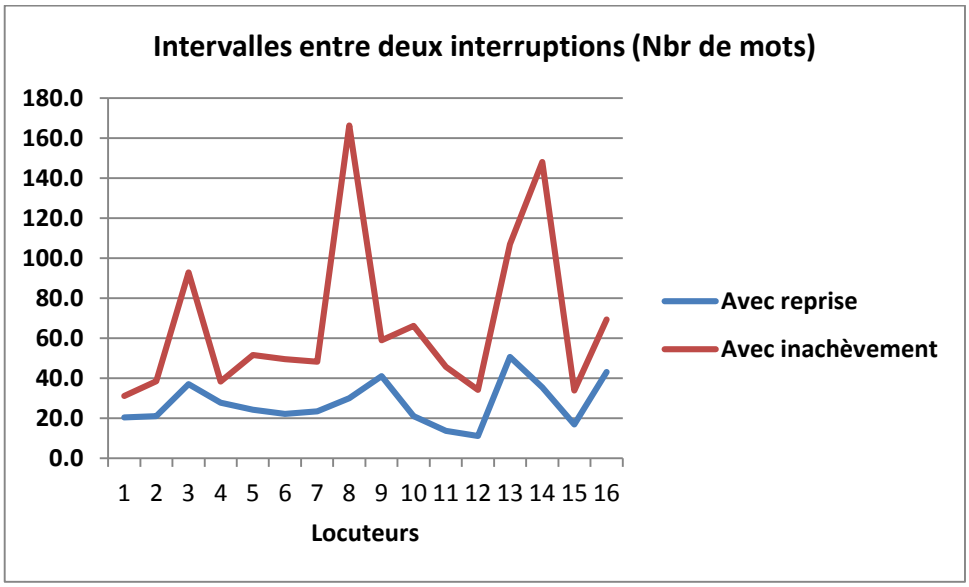



Fig.1 Intervalles entre deux interruptions (en nombre de mots) : comparaison entre disfluences avec reprise et inachèvements de l'énoncé

Si la présence d'un espace Interregnum rempli (pause silencieuse ou non, insertion parenthétique) permet de localiser une rupture dans la fluence verbale, il est des interruptions qui ont lieu sans que cet intervalle soit occupé (il est donc vide). Parmi les interruptions, les disfluentes sont les seules où se produit une reprise de l'énoncé sans que l'intervalle Interregnum soit occupé :

YM_1124 chaque fois **tu ma-//tu *devais* ma**rquer

CM 996 *tu as **des des** feux d'artifices comme ça*

Ces cas sont nombreux puisqu'ils représentent un cinquième (20,4%) de toutes les interruptions et presque la moitié (45,9 %) des interruptions disfluentes. Ils ne sont donc détectables qu'à l'aide des perturbations sémantiques et syntaxiques qui les accompagnent. Cette catégorie d'interruptions disfluentes peut être détectée à l'aide d'un analyseur syntaxique dont les résultats révèleraient automatiquement la présence potentielle de ces perturbations.

**Conclusion**

L'histoire des langues régionales en France et celle de la langue des signes est aussi celle d'une résistance plus ou moins réussie à leur disparition. Cette histoire est liée à l'évolution des conceptions sociétales et psychologiques (et même des moyens technologiques dans le cas de la LSF) à l'égard du bilinguisme qui semble actuellement encouragé.

La constitution de bases de données orales en français parlé s'est heurtée, dès l'origine, à la langue écrite comme norme du bien parler. La valorisation d'un français pur renvoyait le français oral dans le territoire des impuretés de tous ordres et, par là, a freiné l'établissement de grands corpus oraux. L'abandon de la notion de faute au profit de la notion de variante a permis le développement des études sur les caractéristiques de l'oral à partir de gros corpus transcrits et annotés.

Les disfluences dans les énoncés oraux spontanés constituent également un cas de transgression. La langue s'impose de façon particulière comme l'a dit Roland Barthes en 1977 lors de sa leçon inaugurale au Collège de France[10] et c'est tout particulièrement le cas dans la parole spontanée : « La langue, comme performance de tout langage, n'est ni réactionnaire ni progressiste ; elle est tout simplement **fasciste** ; car le fascisme, ce n'est pas d'empêcher de dire, c'est **d'obliger à dire »**.

La transcription des interactions verbales entre locuteurs révèle que l'élaboration de l'oral n'obéit pas sans dysfonctionnements aux contraintes du système de la langue. Pourtant, auto-interruptions, réparations et abandons se succèdent, le plus souvent, sans que l'interaction échoue à se poursuivre. En fait, ces fréquentes transgressions dans la fluidité verbale sont peu perçues et permettent, par la répétition et les « piétinements suspensifs et discursifs », d'alléger l'activité cognitive des locuteurs en présence.

**Bibliographie**

Arnoux, Danièle. « A qui la faute ? ». e.l.p. *Le transfert dans tous ses errata*. Paris : E.P.E.L, 1991. 223-238.

Barkat-Defradas, Mélissa, Coralie Busseuil, Oriane Chauvy, Fabrice Hirsch, Camille Fauth, Joana Révis et Benoit Amy de la Bretèque. « Dimension esthétique des voix normales et dysphoniques : approches perceptive et acoustique ». *Travaux interdisciplinaires sur la parole et le langage* (2012). http://tipa.revues.org/204.

Bartolotti, James et Viorica Marian. « Language Learning and Control in Monolinguals and Bilinguals ». *Cognitive Science* 36 : 6 *(*2012) : 1129-1147.

Bergougnioux, Gabriel. « Etude sociolinguistique sur Orléans (1966-1970) ». *Revue française de linguistique appliquée* 1 : 2 (1996) : 87-88.

Bertrand, Roxane, Philippe Blache, Robert Espesser, Gaëlle Ferré, Christine Meunier, Béatriceù Priego-Valverde et Stéphane Rauzy. « Le CID - Corpus of Interactional Data - Annotation et exploitation multimodale de parole conversationnelle ». *Traitement automatique des langues* 49 : 3 (2008) : 105-134.

Bigi, Brigitte, Roxane Bertrand et Matilde Guardiola. « Recherche automatique d'hétéro-répétitions dans un dialogue oral spontané ». *XVIIIèmes Journées d'études sur la Parole, mai 2010, Mons, Belgium*, 2013. 1-4.

Bilger, Mireille et Claire Blanche-Benveniste. « Français parlé - oral spontané. Quelques réflexions ». (2001). http://icar.univ-lyon2.fr/ecole_thematique/contaci/documents/bilger_cappeau/CBB-Bilger.pdf.

Blanche-Benveniste, Claire et Colette Jeanjean. *Le français parlé. Edition et transcription.* Paris : Didier Erudition, 1986.

---, Mireille Bilger, Christine Rouget et Karl van den Eynde. *Le français parlé. Etudes grammaticales.* Paris : CNRS, 1990.

---, Christine Rouget, Frédéric Sabio, Frédéric, dir. *Choix de textes de français parlé. 36 extraits*. Paris : Champion, 2002.

B.O. hors-série 8. 30 août 2007. Encart paru au B.O 26. 26 juin 2012 : lettre de rentrée 2012 (alinéa 7) sur la continuité de l'enseignement des langues vivantes de l'école au collège.

Boysson-Bardies, Bénédicte de. *Comment la parole vient aux enfants. De la naissance jusqu'à deux ans*. Paris : Odile Jacob, 1996.

Brunot, Ferdinand. *Les archives de la parole.* 1912-1913. http//galica.bnf/ArchivesParole.

Busquet, Denise et Cécile Allaire. *La surdité de l'enfant. Guide pratique à l'usage des parents. Un partenariat entre la Fondation de France et l'Institut national de prévention et d'éducation pour la santé INPES***,** 2012. http://www.inpes.sante.fr/CFESBases/catalogue/pdf/884.pdf.

Byers-Heinlein, Krista, Christopher Fennell et Janet Werker. « The Development of Associative Word Learning in Monolingual and Bilingual Infants ». *Bilingualism : Language and Cognition* 16 : 1 (2013) : 198-205.

Cappeau, Paul et Françoise Gadet. « Où en sont les corpus sur le français parlé ? ». *Revue française de linguistique appliquée* 12 (2007) : 129-133.

--- et Magali Seijido. *Les corpus oraux en français.* Site de la DGLFLF, 2005.

CCNE. Avis 44 rendu le 1er décembre 1994 par le CCNE, 1994.

Cerquiglini, Bernard. *Rapport au Ministre de l'Education nationale, de la Recherche et de la Technologie, et à la Ministre de la Culture et de la Communication*. Paris, 1999.

Clark, Herbert et Thomas Wasow. « Repeating Words in Spontaneous Speech ». *Cognitive Psychology* 37 (1998) : 201-242.

Cordereix, Pascal. « Les archives de la parole ». *Culture et Recherche* 124 (2011) : 31.

Coyos, Jean-Baptiste. « Le basque ». *Histoire sociale des langues de France*. Dir. Georg Kremnitz avec le concours de Fanch Broudic et le collectif HSLF. Rennes : PU de Rennes, 2013. 427-437.

Dister, Anne. « De la transcription à l'étiquetage morphosyntaxique. Le cas de la banque de données textuelles orales VALIBEL ». U de Louvain. Thèse non publiée, 2007.

Duchêne, Roger, dir. *Madame de Sévigné. Lettres de l'année 1671.* Paris : Gallimard, 2012.

Duez, Danièle. Organisation temporelle de la parole et dysarthrie parkinsonienne. 2014. http://aune.lpl-aix.fr/~fulltext/2416.pdf

Encrevé, Florence. « Réflexions sur le congrès de Milan et ses conséquences sur la langue des signes française à la fin du XIXème siècle ». *La découverte/ Le mouvement social* 2 : 223 (2008) : 83-98.

Encrevé, Pierre. « En quoi l'idéologie linguistique française s'oppose-t-elle à la territorialisation des langues ». *Langues régionales : langues de France, Langue d'Europe*. *Actes de la journée d'études au Centre Pompidou*. Paris : BPi Centre Pompidou, 1999. 23-24.

Eshkol, Taravella, Denis Maurel et Isabelle Tellier, Isabelle. « Annoter ESLO1 par des cascades de transducteurs et par apprentissage automatique ». Journée ATALA : Annoter les corpus oraux. Paris, 9 avril 2011.

Gadet, Françoise. *Le français ordinaire.* Paris : Colin, 1997.

Gougenheim, Gabriel, René Michea, Paul Rivenc et Aurélien Sauvageot. *L'élaboration du français fondamental (1° degré). Etude sur l'établissement d'un vocabulaire et d'une grammaire de base.* Paris : Didier, 1964 (rééd. *L'élaboration du français élémentaire*, 1956).

Hagège, Claude. *L'enfant aux deux langues.* Paris : Odile Jacob, 1996.

---. *Combat pour le français. La diversité des langues et des cultures.* Paris : Odile Jacob, 2006.

Hermon-Belot, Rita. *L'abbé Grégoire : la politique et la vérité*. Paris : Seuil, 2000.

INPES. *La surdité de l'enfant. Guide pratique à l'usage des parents. Un partenariat entre la Fondation de France et l'Institut national de prévention et d'éducation pour la santé INPES*. [http://www.inpes.sante.fr/CFESBases/catalogue/pdf/884.pdf].

Julia, Dominique, Michel de Certeau et Jacques Revel. « Une ethnographie de la langue : l'enquête de Grégoire sur les patois ». *Annales. Économies, Sociétés, Civilisations* 30 : 1 (1975) : 3-41.

Lebrave, Jean-Louis. « L'édition génétique ». *Les manuscrits des écrivains*. Paris : CNRS / Hachette, 1993. 206-223.

Marian, Viorika et Anthony Shook. « The Cognitive Benefits of Being Bilingual ». *Cerebrum* 31 (2012) : 1-12.

Morvan, Françoise. *Le Monde comme si - Nationalisme et dérive identitaire en Bretagne*. Actes Sud, 2002.

Muzerelle, Denis. « Aperçu sommaire (et perspectives nouvelles) sur les notes tironiennes ». *Écritures abrégées (notes, notules, messages, codes...) : l'abréviation entre pratiques spontanées, codifications, modernité et histoire.* Dir. N. Andrieux-Rey, S. Branca-Rosoff, C. Puech. Paris / Gap : Ophrys, 2004. 191-210.

Pallaud, Berthille et Sandrine Henry. « Les effets langagiers des amorces de mots et des répétitions dans les énoncés de français parlé spontané ». *Perturbations et réajustements. Langue et langage.* Dir. B. Vaxelaire, R. Sock, G. Kleiber et F. Marsac. Strasbourg : Publications de l'U Marc Bloch Strasbourg 2, 2007. 131-141.

---, Stéphane Rauzy et Philippe Blache. « Auto-interruptions et disfluences en français parlé dans quatre corpus du CID ». *Travaux interdisciplinaires du Laboratoire parole et langage d'Aix-en-Provence* 29 (2013) : 1-19.

Philipp, Marthe. « Annexe de l'article Alsace et Moselle ». *Les langues de France.* Dir. Bernard Cerquiglini. Paris : PUF, 2003. 40-57.

Rivenc, Paul. « Les auteurs du *Français fondamental* face à un objet nouveau et insolite : l'interaction orale ». *Documents pour l'histoire du français langue étrangère ou seconde* 36 (2006) : 35-49.

Shriberg, Elizabeth. « Preliminaries to a Theory of Speech Disfluencies ». U of California at Berkeley. Ph.D Thesis, 1994.

Starkweather, C. Woodruf et Janet Givens-Ackerman. *Stuttering.* Austin, TX : Pro-Ed, 1997.

Stefanini, Jean. « Approches historiques de la langue parlée ». *Recherches en syntaxe sur le français parlé* 5 (1983) : 7-22.

Tabouret-Keller, Andrée et Frédéric Luckel. « Maintien de l'alsacien et adoption du français. Eléments de la situation linguistique en milieu rural en Alsace ». *Langages* 61 (1981) : 39-62.

**---.** « La nocivité mentale du bilinguisme : cent ans d'errance ». *Congrès de la langue basque*. (San Sebastian 1987*)*. Tome III. Vitoria-Gasteiz : Servicio Central de Publicaziones del Gobierno Vasco, 1988. 155-169.

---. *Le bilinguisme en procès.* Limoges : Lambert-Lucas, 2011.

Tardieu, Jean. *La comédie du langage.* Paris : Gallimard, 1987 (1966).

Walter, Henriette. *Aventures et mésaventures des langues de France.* Nantes : Temps, 2008.

Werker, Jane, Nùria Sebastian-Galles, Bàrbara Albareda, et Whitney Weikum. « A Bilingual Advantage in Visual Language Discrimination in Infancy ». *Psychological Science* 23 : 9 (2012).

**NOTES**

[1] Il inclut dans cette liste les 23 langues régionales parlées en Nouvelle-Calédonie.

[2] 21 ans de recul en France : les premières implantations ont eu lieu à Montpellier et Toulouse en 1992.

[3]*Le traitement de la surdité par implants cochléaires ou du tronc cérébral.* Haute Autorité de la Santé. http://www.has-ante.fr/portail/upload/docs/application/pdf/fiche_bon_usage_implants_cochleaires.pdf

[4] L'écriture conserve également la parole mais seulement dans sa modalité textuelle. La transcription textuelle des paroles, moins dépendante de la technique, semble avoir été utilisée dès le IIIème siècle avant notre ère par le grec Xenophon. Le système de notes tironiennes pour transcrire les discours de certains orateurs dont Ciceron (vers 63 de notre ère) a été employé jusqu'au XII siècle (Muzerelle, 2004, 2006).

[5] *http//gallica.bnf.fr*

[6] Cette liste comprend : 380 substantifs, 200 verbes, 100 adjectifs qualificatifs et 50 mots de diverses sortes, soit un total d'un millier.

[7] Par exemple le corpus du CID qui a connu deux campagnes d'enregistrement en chambre sourde : 2003 et 2005 (Bertrand et al., 2008). Ce corpus est constitué de 8 dilogues d'une heure chacun ; il s'agit de conversations entre deux locuteurs filmés et enregistrés sur une piste chacun. Ce corpus fournit des informations sur le signal sonore, la structure syntaxique et les gestes. Ces enregistrements ont été transcrits, alignés et annotés sous Praat. L'annotation multimodale du CID inclut la phonétique, la prosodie, la morphologie, la syntaxe, les disfluences, le discours et la mimo-gestualité. La transcription est orthographique et enrichie de la notation de certains phénomènes de prononciation qui permettront la transcription phonétique des enregistrements ainsi que l'alignement son-token (mot) et graphème-phonème)

**[8]** Trouble acquis de l'émission du langage, qui réalise une répétition incoercible, parfois itérative, de syllabes, de mots, de courtes phrases, et comporte une tachyphémie avec fréquente aphonie terminale dans sa forme majeure.

[9] Le terme « dilogue » a été employé pour désigner les interactions enregistrées et filmées du CID.

[10] http://egophelia.free.fr/pouvoir/barthes.htm