

Troncations de mots, reprises et interruption syntaxique en français parlé spontané

Berthille Pallaud

Parole et Langage, UMR CNRS 6057, Université Aix-en-Provence, 13621, France

Abstract

A specific study of the fragments of words in spoken French were undertaken some years ago on the basis of corpora of non directive talks recorded and transcribed according to Aix Group of Research in Syntax (DELIC currently)'s conventions'. The objective of our study is to describe the types of restarts which follow these interruptions and their localization. Some parameters like the category of truncation, the type of words concerned by the interruption and the site of the fragment of words compared to the verb modulate the presence of the restarts after truncation. It appears that the left unfinished fragments of words are, with some exceptions, the only one, and once on two, followed by a syntactic rupture. These cases of rupture in the verbal construction are undoubtedly which, on the perceptive level, causes this impression of failures in the statement. Our analyses on the interruptions of left unfinished words and followed by a rupture of verbal construction, make it possible to conclude that two parameters seem to influence the way the statement will evolve: there is more chance than the speaker does not continue his proposal, on the one hand, if he interrupts a word before or on the verb and, on the other hand, if this element is a function word.

Résumé

Une étude spécifique des amorces de mots en français parlé a été entreprise depuis quelques années sur la base de corpus d'entretiens non directifs enregistrés et transcrits selon des conventions du Groupe Aixois de Recherche en Syntaxe (DELIC actuellement). L'objectif de notre étude est de décrire les types de reprises qui suivent ces interruptions ainsi que leur localisation. Certains paramètres comme la catégorie de troncation, le type de mots touchés par l'interruption et l'emplacement du fragment de mots par rapport au verbe recteur modulent la présence des reprises après troncation. Il apparaît que les amorces de mots laissées inachevées sont, à quelques exceptions, près les seules, et une fois sur deux, suivies d'une rupture de construction verbale. Ce sont sans doute ces cas de rupture de la construction verbale qui sur le plan perceptif provoque cette impression de ratés dans l'énoncé. De nos analyses sur les interruptions de mots laissés inachevés et suivies d'une rupture de construction verbale, il semble possible de conclure que deux paramètres semblent influencer le devenir de l'énoncé : il y a plus de chance que le locuteur ne poursuive pas sa proposition, d'une part, si il s'interrompt sur un élément avant ou sur le verbe et, d'autre part, si cet élément est un mot-outil.

Mots-clés : fragments of word, repetitions, spoken French, oral corpus

1. Introduction

Si la fluidité d'un énoncé oral se mesure à la régularité rythmique dans sa production, il est clair que les énoncés oraux ne sont pas fluides mais disfluents (Shriberg, 1999 ; Padeloup, 1992). Tout locuteur produit de la parole avec une certaine variabilité dans le débit des mots, des pauses (silencieuses ou non, Duez, 2001), des allongements d'éléments linguistiques, etc. De façon générale, ces phénomènes, peu perçus par les locuteurs en présence (y compris par celui qui parle), déclenchent très rarement des commentaires. Certaines de ces disfluences dans la parole, au demeurant très fréquentes, se caractérisent par une interruption (IP) dans l'énoncé que ce soit au niveau morphologique (l'amorce de mot) ou à la frontière de mots (ce

qui est le cas du phénomène de bribe suivi ou non de répétition de mots ; Schriberg, 1999 ; Henry, 2002) :

ex. 1 Amorce de mot : Arborign, 5, 14 c'est vrai que c'est pas **b**-(IP) **beau** d'associer les deux

ex. 2 Bribe suivie de répétition : Laurent, 1, 2 ils auront leur propre (IP) leur propre langage

Les amorces de mots ont été considérées, le plus souvent, comme des répétitions partielles et leur étude s'est faite dans le cadre des répétitions à l'oral (Levelt, 1989 ; Candea, 2000). Cependant, une analyse comparative de ces deux sortes de disfluecence montre qu'il est préférable de ne pas les confondre (Henry & Pallaud, 2003, 2004 ; Pallaud et Henry, 2004). Outre que les fragments de mots sont deux fois moins fréquents que les répétitions, ils ne font l'objet d'une reprise que dans la moitié des cas (ce qui aboutit à compléter le mot amorcé). Les amorces de mots ne sont donc des répétitions partielles que dans un cas sur deux. D'autres différences sont encore constatées : si les fragments de mots sont le plus souvent des mots lexicaux, les répétitions, elles, concernent au contraire surtout les mots-outils.

Dès l'origine, ces achoppements conçus comme des trébuchements, ont été considérés comme des erreurs et qualifiés de *disfluents* voire même de *dysfluents*. Or, leur étude (en particulier, grâce au dispositif de la mise en grille ; Blanche-Benveniste, 1997), a moins insisté sur l'aspect disfluent de ces phénomènes que souligné la progression de l'énoncé par piétinements syntaxiques successifs (Jeanjean, 1984). Ainsi, ces « haltes » dans la production du texte sont obtenues par les pauses silencieuses ou remplies (nommées pauses d'hésitation, Duez, 1997) mais aussi par les répétitions et une bonne partie des amorces de mots.

L'objectif de cette étude sur les fragments de mots dans les énoncés oraux est de préciser leur localisation syntaxique et le lieu de la reprise de l'énoncé quand il y en a une. Shriberg et Stolcke (1998) ont montré sur les répétitions (où, par définition, il y a toujours une reprise de l'énoncé) que le locuteur tend à remonter au début du constituant qu'il a du mal à reformuler. C'est une question qui sera envisagée à propos de l'amorce seule, du point de vue de sa position (le fragment et sa reprise) et de certains de ses aspects morpho-syntaxiques et syntaxiques.

2. Corpus et méthodologie

L'étude porte sur un ensemble de 20 corpus recueillis (105 000 mots), à l'exception de l'un d'entre eux, dans le cadre d'une enquête menée en 1998 et 1999, par le Groupe Aixois de Recherche en Syntaxe (sous la responsabilité de Claire Blanche-Benveniste) sur le français parlé dans toutes les régions de France. La totalité de cette enquête a été poursuivie et numérisée (son et transcription) par l'équipe DELIC (*DEscription Linguistique Informatisée sur Corpus* : Dir J. Veronis). Selon les conventions du Groupe Aixois de Recherche en Syntaxe qui prévoient une transcription orthographique des énoncés oraux, les amorces de mots sont notées par un trait d'union collé au fragment du mot et donc identifiable automatiquement de façon univoque.

Tous ces enregistrements sauf deux ont été conduits en privé, selon la méthode de l'interview la moins directive possible, et ne rassemblaient que deux locuteurs (l'interviewer et l'interviewé). Les deux corpus faits en public étaient des exposés improvisés devant un groupe de 40 personnes. Tous les locuteurs sont adultes.

Ces corpus ne sont pas étiquetés. Afin de pouvoir analyser les énoncés du seul locuteur interviewé (et non ceux du locuteur interviewer), ils ont été extraits des 20 corpus-sources pour constituer un sous-ensemble de 20 corpus « interviewé ». Chaque corpus correspond

alors à un seul locuteur. C'est dans ce sous-ensemble qu'ont été recherchées les amorces de mots, à l'aide du logiciel Contextes (établi par J. Veronis). Ce logiciel a permis de recenser les amorces de mots présentes, entourées d'un contexte antérieur et postérieur de 10 à 30 mots chacune. Les énoncés contenant ces fragments de mots sont au nombre de 441.

L'analyse porte, donc, exclusivement sur les énoncés du locuteur interviewé. L'étiquetage (non automatique) est fait, dans un deuxième temps, sur les 441 énoncés extraits de ces corpus « interviewé » et rassemblés dans un tableur (Excel). Si on se fonde sur un débit moyen de 200 mots/min, le nombre total de mots étant de 105 000 mots, la durée totale de l'ensemble de ces corpus est de 7 heures 51 min. La longueur moyenne des corpus est de 3080 mots soit une durée moyenne de 16 min. ; les valeurs extrêmes de ces durées sont 1307 et 4931 mots. On retrouve, dans ce sous-ensemble de corpus, la fréquence moyenne d'apparition des amorces de mots dans un énoncé (1 amorce/57sec. ; variation individuelle de 1/23 sec. à 1/8 min.).

Pour chaque énoncé, plusieurs paramètres sont distingués et leurs interactions étudiées. Le premier paramètre concerne le type d'amorce de mot et a été étudié dès le début des travaux sur ce type de phénomène (Pallaud, 2002). Il a été possible de distinguer trois sortes d'interruptions lexicales selon que le fragment se trouve complété, modifié ou laissé inachevé :

A. complétée : Tropr102, 3,3 on va **com-** on va **commencer** à les + à le houspiller un peu

A. modifiée : Tropr101, 1, 3 euh et c'est vrai en plus on (n') a **jam-** on (n') a **pas** compris

A. inachevée : Pariscen, 1, 5 alors il y a les O.N.G. + et puis les les **gouver-** les qui qui qui donnent des livres gratuitement

Pour cette étude les paramètres qui ont été recherchés sont :

la position syntaxique du fragment et sa localisation dans le syntagme ;

la localisation de la reprise, quand il y en a une, après l'interruption lexicale ;

la poursuite ou l'interruption de l'énoncé après la reprise.

3. Les résultats : analyse des réajustements après les interruptions de mots

On constate que deux sortes de phénomènes de réajustement après les amorces de mots peuvent être décrits : d'une part, les **insertions** d'éléments linguistiques, d'autre part les **reprises** de l'énoncé (que ce soit par piétinement syntaxique ou par la poursuite de l'énoncé sur l'axe syntagmatique). Nous emploierons le terme de « reprise » uniquement dans le cas où le locuteur piétine sur la même place syntaxique c'est à dire reprend une partie de son énoncé avant de la poursuivre (ce qui correspond aux cas des amorces complétées ou modifiées). Cette étude ne décrira que les phénomènes de reprise.

3.1. Les conditions de la reprise

Le résultat massif sur le devenir des amorces est que lorsque le fragment de mot est complété, il l'est toujours (et cela sans aucune exception dans nos corpus) à travers au minimum une reprise du fragment. Il n'y a jamais de simple complétude de fragment par le fragment manquant (du type : un **li- vre**). C'est un trait « qualitatif » qui permet de distinguer un énoncé disfluent d'un énoncé produit par une personne atteinte de « béguayage » (Zellner, 1992).

Par définition, la reprise après interruption de mot n'a lieu que dans les cas d'amorces complétées ou modifiées qui sont les seuls fragments de mots suivis d'un piétinement sur la même place syntaxique. Les amorces inachevées, elles, sont justement identifiées parce que ce qui suit le fragment n'appartient pas à la même place syntaxique que lui. Le réajustement, dans ce cas-là, n'est pas une reprise mais une poursuite de l'énoncé.

Les proportions de ces trois catégories d'amorces vérifient les conclusions de Levelt (1989) sur les phénomènes d'interruption en général dans les énoncés : ces phénomènes s'accompagnent le plus souvent de reprises plutôt que d'interruption définitive de la proposition. Dans nos corpus, en effet, les locuteurs reprennent (amorces complétées et modifiées ; 78%) plus qu'ils ne poursuivent leur énoncé sans revenir en arrière (amorces laissées inachevées ; 22%) après une interruption. Restait à étudier comment se fait cette reprise de l'énoncé après une interruption de mot.

Nous nous sommes intéressés à la position de l'élément interrompu et à la localisation de la reprise qui suit puisqu'on ne peut décrire la localisation des reprises après interruption de mot sans préciser auparavant la **localisation de cette interruption** qui peut avoir lieu en tête ou non du groupe nominal ou verbal :

- Fragment **En tête** :

ex. 3 C5bBelfo 1 parce qu'il y a **u- une** frange de jeunes qui arrivent maintenant

- Fragment **Non en tête** :

ex. 4 C7dBord 5,5 et et après j'ai **vou- voulu** changer

ex. 5 C24aNanc3 22, 3 donc les les **ani-** les **animaux** et les coraux qu'on trouve là

Si l'amorce de mot est déjà en tête de groupe (Syntagme nominal : **u-une frange**), la localisation en tête de la reprise n'a pas la même signification que si l'amorce ne s'y trouve pas déjà (Syntagme nominal : reprise en tête, *les ani- les animaux* ; Reprise non en tête, *les ani- animaux*). En effet, si le fragment de mot n'est pas en tête de groupe mais que la reprise remonte jusqu'à cette tête, cela traduit une contrainte linguistique qu'il n'est pas possible de mettre en évidence lorsque l'amorce est déjà en début de groupe ; c'est pourquoi il nous a paru utile de distinguer ces deux sortes de localisation de l'amorce de mot.

3.1.1 Localisation du fragment de mot

Sur les 436 interruptions de mots où il a été possible de déterminer si elles avaient lieu en tête de groupe ou non (Fig.1), plus des deux tiers (72%) ne se produisent pas en tête de groupe (nominal ou verbal ; pas de différence). Ce résultat n'est pas surprenant puisque la grande majorité de ces disfluences touchent, on l'a vu le lexique, item qui en français est rarement en tête de groupe mais précédé d'un déterminant ou d'un pronom.

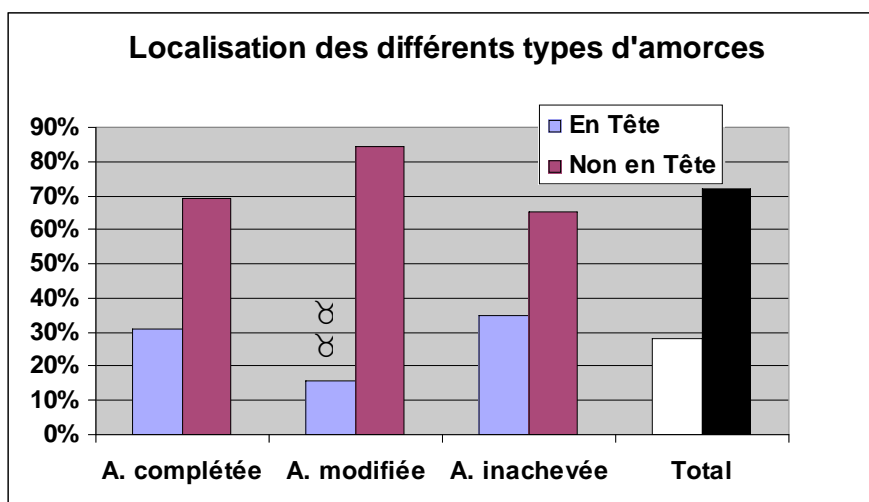


Figure 1 : Localisation (en tête ou non de syntagme) des différents types d'amorces de mots

Les amorces complétées et inachevées ont une répartition semblable dans la localisation au sein du syntagme : 31% et 35% d'entre elles sont en tête (différence non significative).

Seule la répartition des fragments qui seront modifiés diffère significativement¹. Ces amorces sont moins nombreuses lorsque la perturbation n'est pas située en tête de groupe (16%). Ce résultat tient sans doute au fait que c'est l'amorce lexicale et non le déterminant ou le pronom qui est plus souvent remplacée (le nom ou le verbe se trouvant rarement en tête de syntagme).

Les reprises après troncation de mot sont-elles influencées par la localisation du fragment de mot ?

3.1.2. Emplacement de la reprise suivant un fragment de mot qui n'est pas en tête de groupe.

Les reprises observées à la suite des amorces non en tête de groupe peuvent être regroupées dans trois catégories :

*Les reprises « **minimales** » qui ne remontent pas jusqu'au début du groupe :

ex. 6 CorpusEDF, 12,6 la note est est salée hein la note **énergé- énergétique**

*Les reprises « **Début du Groupe** » qui remontent jusqu'au début du Groupe nominal ou verbal :

ex. 7 C5cBelfo, 2,2 ce monsieur s'est approché et il a dit **vous li- vous lisez** l'Est Républicain

*Les reprises « **Plus** » qui remontent au delà du groupe :

ex. 8 Corpus EDF, 8,38 mais l'eau **est de vingt mè- l'entrant est de vingt mètres** cubes

Les reprises après troncation de mot ont lieu le plus souvent au début du syntagme (63% des cas). Ce type de disflunce confirme les observations faites par Blanche-Benveniste (1997 p 48) et les résultats signalés sur tous les types de disflunce par Clark et Wasow (1998) : après une interruption, le locuteur reprend généralement depuis le début d'un syntagme. Il reste qu'une part non négligeable des reprises ne se conforme pas à ce modèle puisque **37%** d'entre elles ne se font pas au début du groupe nominal ou verbal. Certaines, les reprises

¹ (khi2 = 10,46 ; p<.01 ; d.d.l.=2)

« minimales » (24%) ne remontent pas au début du syntagme et d'autres, les reprises « Plus » (13%), au contraire, dépassent cette limite.

3.1.3. Emplacement de la reprise qui suit un fragment de mot situé en tête de groupe.

Dans nos corpus, seuls 28% des amorces de mots sont en tête de groupe (Fig.1). Comment s'effectue la reprise dans ces cas-là ? Va-t-on trouver des reprises « Plus » ? Les amorces complétées et modifiées (qui sont donc suivies d'une reprise) sont, dans leur grande majorité (92%), suivies d'un réajustement localisé sur la place du segment de mot, donc en tête de groupe.

ex. 9 C5cBelfo 16 ,1 j'ai d'abord été assez suffoquée j'ai eu **div- diverses** réactions

Rien ne s'oppose cependant à ce que la reprise ait lieu plus avant dans l'énoncé (reprise « Plus »). Quelques unes (6%), en effet, dépassent cette limite (/2) et remontent sur des éléments antérieurs (/1) :

ex. 10 C6cBesan, 6, euh dans la mise en pratique [1je vois pas [2c- [1je verrais pas /2**comment** faire correspondre les choses

ex. 11 Tropr102, 3,1, [1je crois que [2c- + [1je crois que [2c'**est** la famille des N. mais j'en suis pas sûr

On constate, donc, que le locuteur sort, aussi souvent, des limites du syntagme lorsqu'il reprend plus avant son énoncé (reprise « Plus ») que l'interruption ait lieu en tête ou non du syntagme. Ce type de reprise ne se produit que dans 10% des cas environ et ne dépend pas de la localisation de l'amorce de mot.

3.1.4. Emplacement de la reprise et types d'amorces

Le type d'amorce est un paramètre qui exerce une influence sur la localisation de la reprise qui suit. Le locuteur qui reprend son énoncé après une interruption de mot ne le fait pas de la même façon si le fragment est complété, modifié ou laissé inachevé (Fig.2). Pour compléter son amorce, le locuteur reprend significativement plus souvent² son énoncé au début du syntagme que lorsqu'il le modifie. Dans ce dernier cas, il augmente la fréquence de deux stratégies : reprise minimale et reprise d'éléments antérieurs au syntagme concerné par l'interruption (Reprise « Plus »)

3.2. Troncations et rupture syntaxique

Nous avons vu que la troncation peut s'accompagner d'un piétinement syntaxique (amorces complétées et modifiées) ou non (amorces inachevées). Le plus souvent, même lorsqu'il s'agit d'un segment qui sera laissé inachevé, cette interruption ne s'oppose pas à la poursuite de l'énoncé c'est à dire à une continuité syntaxique. Le locuteur continue de développer le syntagme suspendu et produit d'autres arguments pour le même verbe recteur :

ex. 12 C5cBelfo, 2,2 ce monsieur s'est approché et il a dit vous **li-** vous **lisez** l'Est Républicain donc vous êtes de l'Est

Dans d'autres cas au contraire, ce qui suit le fragment ou son réajustement correspond à **une rupture dans la construction syntaxique** (c'est à dire que ce qui suit n'appartient plus à la rection du verbe) :

² Khi2=11,64 ; p<.001 ; d.d.l.=2

ex. 13 Pariscentre, 1, 1 euh **c'ét- c'était** ça m'a demandé beaucoup d'effort de compréhension

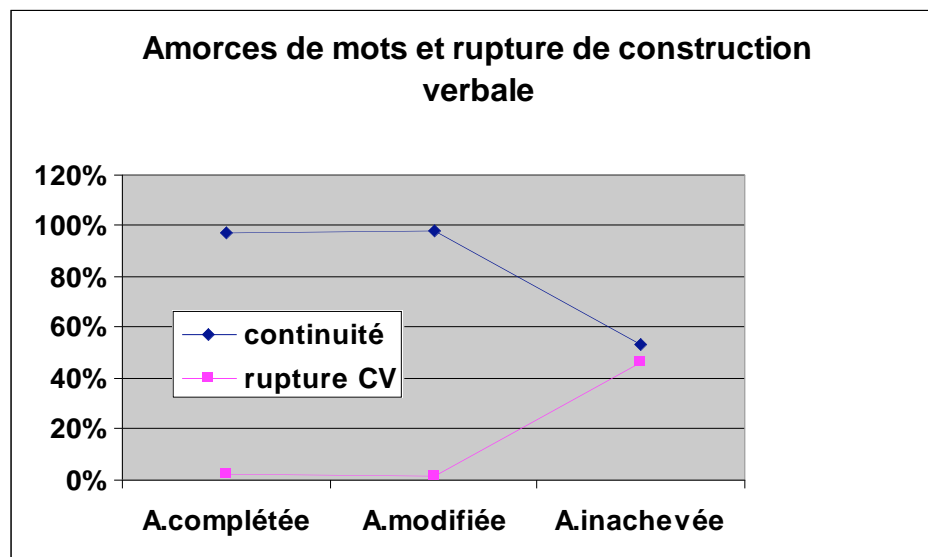


Figure 2. Rupture de construction verbale après des amorces de mots complétées, modifiées et inachevées

On constate que les troncations de mots (toutes confondues) sont, dans la grande majorité des cas (88%) suivies d'une poursuite de la proposition. Cette interruption ne provoque que dans 12% des cas seulement une rupture de construction verbale. Cependant, si des exemples de ces cas sont observés dans les trois types de fragments de mots, ils le sont dans des proportions considérablement différentes (Fig. 2)³. À quelques exceptions près, la majorité des amorces de mots (qui seront complétées ou modifiées : 78% des amorces de mots donc) ne sont pas suivies d'une rupture syntaxique. Il n'en est pas de même pour l'amorce qui est laissée inachevée (22%) : près de la moitié (46%) de ces troncations sont suivies d'une rupture de construction verbale.

La localisation de l'amorce de mot par rapport au verbe recteur a également une influence sur l'impossibilité pour le locuteur de maintenir la continuité de son énoncé. Les cas de rupture de construction verbale sont deux fois plus nombreux avant et sur le verbe qu'après le verbe. Autrement dit, plus la troncation a lieu tôt (avant et sur le verbe), plus l'énoncé risque de s'interrompre⁴. En revanche, les cas de continuité syntaxique sont aussi nombreux, quelle que soit la localisation du fragment de mot.

Comme ces ruptures sont réservées presque aux seuls cas d'interruptions laissées inachevées, leur plus grande fréquence sur la zone avant/sur le verbe pourrait être due au fait que les amorces laissées inachevées se produisent plus souvent dans cette zone. Nous vérifions qu'il n'en est rien. Lorsque l'énoncé se poursuit sans rupture syntaxique, la répartition des amorces inachevées par rapport au verbe est la suivante : 60% après le verbe et 40% avant ou sur le verbe. Lorsqu'il y a rupture syntaxique, ces proportions sont inverses : 24% des amorces inachevées sont après le verbe et 76% le sont avant ou sur le verbe.

³ $\text{Khi}^2=138,78$; $p<.0001$; d.d.l.=2

⁴ $\text{Khi}^2=6,44$; $p<.02$; d.d.l.=1

Le type de mots semble, aussi, avoir une influence sur la possibilité qu'a le locuteur de pouvoir compléter sa proposition si l'amorce est laissée inachevée. Les mots-outils interrompus et laissés inachevés sont significativement⁵ plus souvent que les mots-pleins suivis d'une rupture de la construction verbale (65% contre 33%).

3.3. Conclusion

Nous avons étudié les réajustements auxquels le locuteur procède après s'être interrompu au milieu d'un mot d'un double point de vue : d'une part, les types d'insertions (30% des cas) et d'autre part les reprises de l'énoncé (75% des cas). Seules les reprises sont abordées ici.

Les interruptions de mots se produisant dans la grande majorité des cas sur du lexique, elles ne sont donc pas en tête de groupe nominal ou verbal. La reprise peut se faire sur le mot même interrompu, remonter jusqu'au début du groupe (nominal ou verbal) ou même dépasser cette limite. Si les cas les plus nombreux sont bien (comme l'avaient signalé Glark et Wasow, 1998) ceux où la reprise de l'énoncé remonte jusqu'au début du groupe, un tiers des reprises échappent à ce schéma, ce qui n'est pas négligeable. Près de 24% des reprises sont en effet « minimales » et ne remontent pas jusqu'au début du groupe. Les autres au contraire dépassent la limite du Groupe nominal ou verbal. On constate que les reprises (6 à 10 %) qui dépassent les limites du groupe où a lieu l'interruption ne dépendent pas de la localisation de cette dernière. En revanche, lorsque le locuteur complète le fragment de mots interrompu, il remonte plus volontiers au début du syntagme qu'il ne le fait lorsqu'il modifie le fragment interrompu. Dans le cas d'amorce modifiée, il adopte plus souvent les stratégies « minimale » ou « Plus » indifféremment.

Contrairement à ce qui se passe pour les amorces complétées et modifiées qui ne provoquent qu'exceptionnellement un arrêt du déroulement de la proposition, dans 46% des cas, ce qui suit le fragment inachevé interrompt la proposition. Ce sont sans doute ces cas de rupture de la construction verbale qui, sur le plan perceptif, provoque cette impression de ratés dans l'énoncé. De nos analyses sur les interruptions de mots laissés inachevés et suivies d'une rupture de construction verbale, il semble possible de conclure que deux paramètres semblent influencer le devenir de l'énoncé : il y a plus de chance que le locuteur ne poursuive pas sa proposition, d'une part, s'il s'interrompt sur un élément avant ou sur le verbe et, d'autre part, si cet élément est un mot-outil.

Références

- Blanche-Benveniste C. (1997). *Approches de la langue parlée en français*. Éditions Ophrys, Paris.
- Candéa M. (2000). *Contribution à l'étude des pauses silencieuses et des phénomènes dits « d'hésitation » en français oral spontané. Étude sur un corpus de récits en classe de français*. Thèse d'État, Université Paris III (Sorbonne Nouvelle). <http://www.anrtheses.com/fr/>
- Clark H. et Wasow T. (1998). Repeating words in spontaneous speech. *Cognitive Psychology*, 37 : 201-242.
- Duez D. (1997). La signification des pauses dans la production et perception de la parole. *Revue Parole*, 3-4 : 275-299.
- Duez D. (2001). Signification des hésitations dans la production et la perception de la parole spontanée, *Revue Parole*, 17-18-19 : 113-138.

⁵ khi2=8,85 ; p<.01 ; d.d.l.=1

- Henry S. & Pallaud B. (2003). Word fragments and repeats in spontaneous spoken French. In Eklund R., *Disfluency in Spontaneous Speech Workshop*, Proceedings of DiSS'03, 5-8 September 2003, Göteborg University, Sweden : 77-80.
- Henry S. & Pallaud B. (2004). Amorces de mots et répétitions dans les énoncés oraux, *Recherche Sur le Français Parlé*, 18 : 201-229.
- Jeanjean C. (1984). " Les ratés c'est fa- fabuleux". Étude syntaxique des amorces et des répétitions. *LINX*, 10, *Syntaxe et discours* : 171-177.
- Levelt W.J.M. (1989). *Speaking. From intention to articulation*. MIT Press, Cambridge.
- Pallaud B. (2002). Les amorces de mots comme faits autonymiques en langage oral, *Recherches Sur le Français Parlé*, 17 : 79-102.
- Pallaud B. & Henry S. (2004). Amorces de mots et répétitions : des hésitations plus que des erreurs en français parlé. In *Le poids des mots. Actes des 7èmes Journées Internationales d'Analyse statistique des Données Textuelles*, Louvain-la-Neuve, 10-12 mars 2004. Louvain, PUL, vol 2 : 848-858.
- Pasdeloup V. (1992). A Prosodic Model for French Text-to-Speech Synthesis : A Psycholinguistic Approach. In Bailly G., Benoit C. et Sawallis T.R. (eds), *Talking Machines. Theories, Models, And Designs* : 335-348
- Shriberg E. & Stolcke A. (1998). How Far Do Speakers Back Up In Repairs? A Quantitative Model. *Proc. Intl. Conf. on Spoken Language Processing*, vol. 5 : 2183-2186.
- Shriberg E. (1999). Phonetic Consequences of Speech Disfluency. Symposium on The Phonetics of Spontaneous Speech (S. Greenberg and P. Keating, organizers), *Proc. International Congress of Phonetic Sciences*, vol. 1 : 619-622.
- Zellner, B. (1992). Le bé- bégayage et euh ... l'hésitation en français spontané. *Actes des 19 èmes Journées d'Études sur la Parole*, J.E.P. Bruxelles : 481-487.

