Réduction temporelle en français spontané : où se cache-t-elle ? Une étude des segments, des mots et séquences de mots fréquemment réduits

Yaru Wu, Martine Adda-Decker

▶ **To cite this version:**

Yaru Wu, Martine Adda-Decker. Réduction temporelle en français spontané : où se cache-t-elle ? Une étude des segments, des mots et séquences de mots fréquemment réduits. 6e conférence conjointe Journées d'Études sur la Parole (JEP, 33e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition). Volume 1 : Journées d'Études sur la Parole, 2020, Nancy, France. pp.627-635. hal-02798596v3

HAL Id: hal-02798596
https://hal.science/hal-02798596v3

Submitted on 23 Jun 2020

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.

Actes de la 6e conférence conjointe Journées d'Études sur la Parole (JEP, 33e édition),
Traitement Automatique des Langues Naturelles (TALN, 27e édition),
Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition)
Nancy, France, 08-19 juin 2020

Réduction temporelle en français spontané : où se cache-t-elle ?

Une étude des segments, des mots et séquences de mots fréquemment réduits

Yaru Wu[1, 2] Martine Adda-Decker[1, 3]
(1) Laboratoire de Phonétique et Phonologie (LPP), UMR7018, CNRS, France
(2) Modèles, Dynamiques, Corpus (MoDyCo), UMR 7114, CNRS, France
(3) LIMSI-CNRS bât. 508, BP 133, 91403 Orsay cedex, France
yaru.wu@sorbonne-nouvelle.fr, madda@limsi.fr

RÉSUMÉ

Cette étude vise à proposer une méthode adaptée à l'étude de divers phénomènes de variation dans les grands corpus utilisant l'alignement automatique de la parole. Cette méthode est appliquée pour étudier la réduction temporelle en français spontané. Nous proposons de qualifier la réduction temporelle comme la réalisation de suites de segments courts consécutifs. Environ 14% du corpus est considéré comme réduit. Les résultats de l'alignement montrent que ces zones impliquent le plus souvent plus d'un mot (81%), et que sinon, la position interne du mot est la plus concernée. Parmi les exemples de suites de mots les plus réduits, on trouve des locutions utilisées comme des marqueurs discursifs.

ABSTRACT

Temporal reduction in spontaneous French : where is it hidden ? A study of frequently reduced segments, words and word sequences

This study aims to propose a method to explore large corpora using automatic speech alignment and to apply this method to the special case of reduction in spontaneous French. By locating sequences of short segments of at least three 30 or 40 ms segments, we were able to identify potential reduction zones. 14% of the corpus is considered as temporally reduced. Short segment sequences are most often observed in cross-word position (81%) rather than in single words. In the latter, the corresponding phone segments are frequently located in word-internal position. The identified reduced sequences often concern phrases used as discourse markers, which carry very little semantic content in real-life communication.

MOTS-CLÉS : français spontané, réduction temporelle, durée segmentale, suite de segments courts, grand corpus.

KEYWORDS: spontaneous French, temporal reduction, segmental duration, sequence of short segments, large corpus.

1 Introduction

La variation de la parole est très présente en parole continue (Duez, 1997; Ernestus, 2000; Duez, 2003; Johnson, 2004; Meunier & Espesser, 2011). Dans la communication quotidienne, les mots sont souvent articulés avec moins de précision et les segments sont affaiblis par rapport à une forme standard. Grâce au traitement automatique de la parole et à la quantité

Volume 1 : Journées d'Études sur la Parole, pages 627–635.
HAL : hal-02798596.
Cette œuvre est mise à disposition sous licence Attribution 4.0 International.

croissante de données accessibles, nous sommes aujourd'hui en mesure d'explorer la parole spontanée et d'étudier des phénomènes variés globalement, sans forcément formuler une hypothèse précise et recueillir des données contrôlées afin d'examiner cette hypothèse en question. Les outils de traitement automatique permettent ainsi des approches différentes de l'analyse traditionnelle basée sur des hypothèses sur un ensemble de données contrôlées.

Un des deux objectifs de cette étude est de suggérer une méthode utilisant l'alignement automatique de la parole afin d'aider à localiser des zones de réduction temporelle. Nous visons également à mieux comprendre où se situent ces zones. Sont-elles principalement situées au sein d'un mot ou vont-t-elle au-delà de la frontière des mots ? Y a-t-il une position dans le mot qui favorise la réduction ? Notre hypothèse de travail ou notre question d'intérêt consiste donc dans le fait que la parole spontanée contienne de nombreux phénomènes de réduction temporelle. Par réduction temporelle nous entendons des zones de parole où les prononciations des mots seraient seulement partiellement réalisées et qu'elles contiendraient potentiellement moins de segments que le nombre prévu par une prononciation canonique. Cette question est intéressante à notre avis non seulement pour les technologies vocales, comme la reconnaissance ou la synthèse de la parole, mais également pour l'apprentissage des langues et pour mieux comprendre le traitement cognitif de la parole.

2 Corpus et alignement

Le corpus NCCFr (Nijmegen Corpus of Casual French, Torreira *et al.*, 2010) a été utilisé dans cette étude. Ce corpus est composé de 36 heures d'enregistrements de conversations entre amis, dont 24 femmes et 22 hommes.

Les données du corpus ont été automatiquement segmentées et étiquetées à l'aide du système de reconnaissance automatique de la parole au LIMSI (Gauvain *et al.*, 2002) en mode d'alignement forcé. Des modèles acoustiques de phones indépendants du contexte (HMM de phones estimés à partir de segments d'au moins 50 ms) ont été utilisés afin d'établir au mieux la correspondance entre les segments de parole et les modèles HMM selon la ou les transcriptions phonémiques proposées par le dictionnaire de prononciation du système. Des variantes de prononciation peuvent être ajoutées à ce dictionnaire pour qu'elles soient mieux adaptées à la production réelle des locuteurs. Le système produit, à la sortie de l'alignement, les étiquettes des mots et des phones, ainsi que les frontières respectives. Le système fournit également des étiquettes et les frontières pour les tronçons de signal hors parole, tels que les pauses, la respiration et le bruit. La durée minimale d'un segment est de 30 ms, ce qui correspond à 3 trames acoustiques (Adda-Decker & Lamel, 2000). Le dictionnaire utilisé contient comme variantes systématiques la liaison et le schwa.

3 Méthodologie

Dans la suite, nous décrivons comment nous proposons de qualifier les zones de réduction et la méthode adoptée afin de les localiser automatiquement dans le signal de parole.

3.1 Méthode ascendante

Nous avons décidé de localiser les phénomènes de réduction temporelle en exploitant les caractéristiques du système d'alignement forcé : ce dernier cherche à associer à chaque phone prévu dans la prononciation d'un mot une portion de signal, dont la durée minimale est au moins de 30 ms. Lorsque, dans le signal observé, certains phones prévus par la prononciation du dictionnaire ne sont pas ou presque pas présents, le système force quand-même l'alignement de ces phones, ce qui engendre, comme résultat de l'étape d'alignement, une séquence de plusieurs segments courts (de 30 ou 40 ms) à la suite. Ce défaut du système d'un point de vue de la précision de l'étiquetage et de la segmentation en phones, peut être exploité comme qualité afin de localiser ces phénomènes de réduction temporelle. Ainsi, nous qualifions comme zone de réduction temporelle une séquence de plusieurs segments courts (30 ou 40 ms) consécutifs produits lors de l'alignement forcé. Ainsi, nous sommes en mesure de séparer les segments dans nos données en deux parties : les segments « normaux » (« Nrm ») et les segments « d'alerte potentiellement réduits » (« Alrt »). Nous ajoutons une contrainte : les segments courts sont étiquetés comme « Alrt », uniquement s'ils sont à l'intérieur d'une séquence d'au moins 3 segments courts consécutifs. Cette méthode, que nous appelons « méthode ascendante », nous permet d'identifier les mots et séquences de mots qui sont réduits d'après ces critères. Par exemple, la séquence de mots « je ne sais pas » (/ʒənəsɛpa/, 4 syllabes en français standard) peut être prononcée [ʃpa] (séquence monosyllabique). Une telle prononciation n'est pas modélisée de manière satisfaisante par le système. L'alignement le mettra en évidence en forçant la présence de tous les segments mais en leur attribuant des durées très courtes.

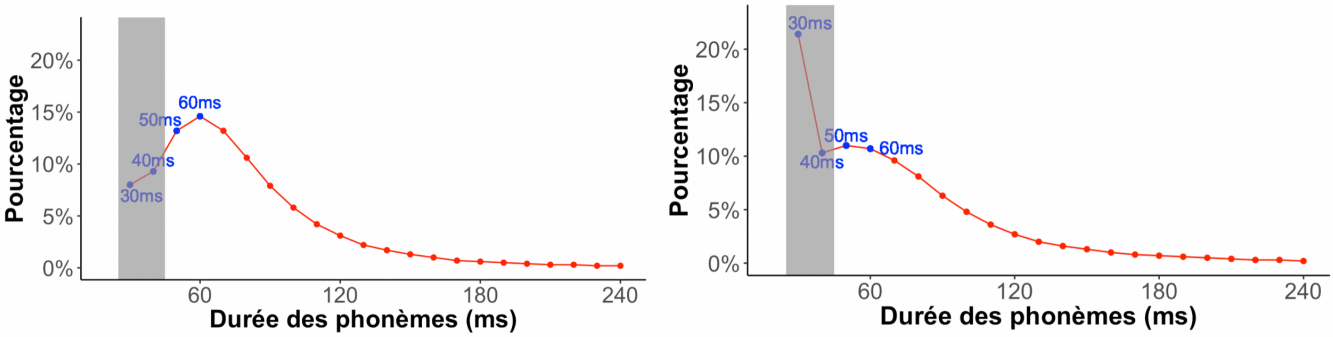

FIGURE 1 – Distribution de la durée des segments (a) dans le corpus journalistique formel ESTER (Galliano *et al.*, 2009) et (b) dans le corpus spontané familier NCCFr (Torreira *et al.*, 2010). L'abscisse concernant la durée segmentale est donnée en milliseconde. L'ordonnée indique le pourcentage de cette durée dans chaque corpus.

La figure 1 illustre la répartition des durées de phones entre (a) la parole formelle journalistique et (b) la parole spontanée familière. La durée des phones sur l'abscisses est indiquée en millisecondes (ms). Les segments potentiellement réduits (à gauche) sont encadrés en gris. La distribution des durées de phones sur la parole formelle journalistique (Figure 1a) est présentée ici pour nous aider à mieux comprendre les caractéristiques de la distribution des durées de phones provenant de parole spontanée familière (Figure 1b). Comme nous pouvons le voir sur la figure 1a, la distribution de la durée des phones

correspond à une courbe globalement en forme de cloche et le sommet de la courbe se situe à 60 ms (voir l'abscisse). Cela suggère que la durée la plus fréquemment observée dans la parole formelle journalistique est de 60 ms. En ce qui concerne la parole spontanée (Figure 1b), le sommet de la courbe correspond à la durée minimale de 30 ms, ce qui suggère que la durée la plus fréquemment observée ici est de 30 ms (>20 %).

Ces observations sont conformes à celles d'Adda-Decker & Lamel (2017) sur la durée des phones en parole préparée et en parole spontanée (en français et en anglais). La parole spontanée familière traitée par l'alignement forcé, de la même manière que la parole formelle, contient alors une proportion extrêmement élevée de segments de durée minimale de 30 ms et de 40 ms. Dans cette étude, nous nous intéressons à la zone en gris de la parole spontanée familière de la Figure 1b.

3.2 Traitement des données

Nous pouvons constater sur la figure 1b que les segments courts (30 ou 40 ms) sont très fréquents dans notre corpus NCCFr de parole conversationnelle familière. Dans la suite, nous expliquerons comment nous avons localisé les suites de segments d'au moins trois segments consécutifs de 30 ou 40 ms.

L'alignement automatique fournit dans la forme de surface résultante (1) les segments d'une durée supérieure à 40 ms, (2) les segments avec moins de trois segments courts (30 ou 40 ms) consécutifs, ou (3) avec au moins trois segments courts (30 ou 40 ms) de suite. (1) et (2) sont considérés comme des segments « normaux » sans rien à signaler (nommés « Nrm ») ; (3) est considéré comme une zone qui est potentiellement impacté par la réduction (nommée « Alrt »). Le tableau 1 donne des exemples sur la façon dont les segments sont classés en fonction de leur durée.

Ex. /stʁ/ du mot « ministre » /ministʁ/	
– Si les segments [s], [t] et [ʁ] (qui se suivent) sont alignés chacun avec une durée courte (30 ou 40ms)	→ [s] segment en alerte : « Alrt » → [t] segment en alerte : « Alrt » → [ʁ] segment en alerte : « Alrt »
– Si les segments [s] et [t] sont alignés chacun avec une durée courte (30 ou 40ms) et le [ʁ] est aligné avec une durée de 50ms	→ [s] segment sans alerte : « Nrm » → [t] segment sans alerte : « Nrm » → [ʁ] segment sans alerte : « Nrm »

TABLE 1 – Exemple illustrant la catégorisation des segments comme « Nrm » ou « Alrt » à partir du mot « ministre ».

4 Résultats

Ci-après, nous étudions les suites de segments courts (zones de réduction potentielle) par rapport aux mots. Est-ce que la zone « Alrt » se trouve à l'intérieur d'un seul mot ou impacte-elle au moins deux mots ? Ensuite, nous regardons dans quelle position se trouvent les segments qui composent les zones « Alrt » à l'intérieur du mot. Ces résultats seront suivis d'une analyse de quelques mots et séquences de mots fréquents.

4.1 Position des segments réduits dans les mots

Nous obtenons un total de 23,002 zones « Alrt » impliquant un total de 86,265 segments. Pour information, le corpus contient un total d'environ 623,844 segments. Nous voulons savoir où se trouvent les séquences de segments étiquetés « Alrt » en opposant la position interne au mot à une localisation chevauchant deux ou plusieurs mots (figure 2). Ensuite, nous regardons pour les segments composant les séquences « Alrt », leur position à l'intérieur d'un mot (figure 3).

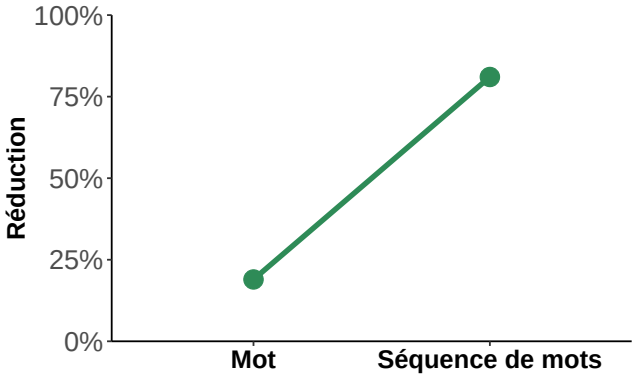

FIGURE 2 – Taux des suites de segments courts dans les mots et les séquences de mots.

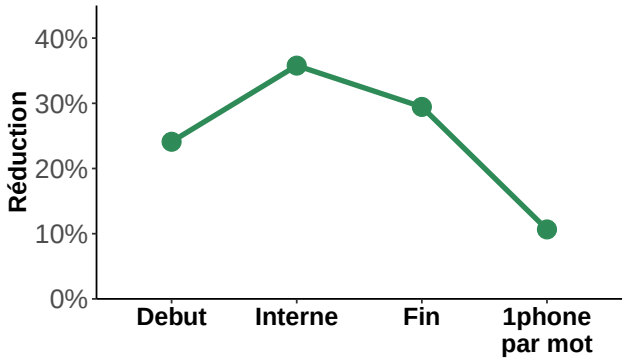

FIGURE 3 – Taux de segments courts en fonction de sa position dans le mot.

La figure 2 montre la proportion de suites de segments courts se réalisant complètement à l'intérieur d'un mot (Mot) [1] et celles impactant au moins deux mots (Séquence de mots) [2] observée dans les zones de réduction. En effet, parmi les « suites de segments courts », la plupart des segments courts font partie des suites de phones qui couvrent plus d'un mot (81%) et les suites de segments courts ne se limitent que rarement à un seul mot (19%). Ces résultats suggèrent que la zone de réduction concerne plus d'un mot en général.

La figure 3, montre la proportion de segments courts en fonction de leur position dans le mot. Nous observons plus de segments en position interne de mot qu'en position finale de mot (et davantage en position finale qu'en initiale) dans la zone de réduction. Cela suggère que la position interne du mot est la position préférée pour la réduction. La proportion la plus faible se trouve sur « 1phone par mot » (c'est-à-dire un segment qui est aussi un mot).

La figure 4 détaille les résultats de la figure 3 en fonction de la nature des suites de segments présentées dans la figure 2 (mot vs. séquence de mots). En ce qui concerne les suites de segments courts qui sont au sein d'un mot, la position interne de mot est la position privilégiée pour la réduction (80%). Il est intéressant de noter qu'aucune tendance particulière n'est observée en ce qui concerne la position dans un mot lorsqu'il s'agit de suites de segments courts au-delà de la frontière des mots (28% vs. 26% vs. 33% pour « Début », « Interne » et « Fin » respectivement) et que la position « 1 phone par mot » a une proportion légèrement inférieure par rapport aux trois autres positions.

1. Ex. suite de segments courts [ʁkw] provenant de la prononciation du mot « pourquoi » /puʁkwa/.
2. Ex. suite de segments courts [aʁʁa] provenant de la séquence de mots « par rapport » /paʁ#ʁapɔʁ/.

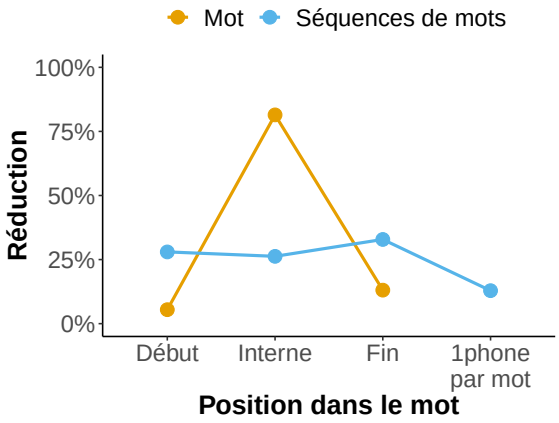

FIGURE 4 – Taux de segments courts suivant la position du segment dans le mot (Début vs. Interne vs. Fin de mot) et la nature des séquences (Mot vs. Séquence de mots).

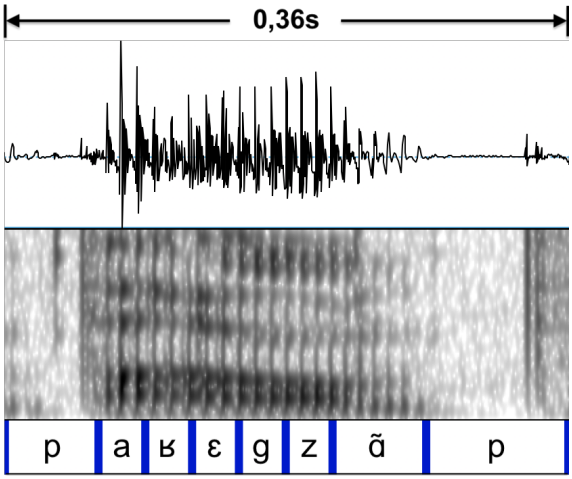

FIGURE 5 – Spectrogramme (0˜5000Hz) et signal de la locution *par exemple* /paʁ#ɛgzɑ̃pl/ (NCCFr, 26_11_07_nb1_2_16.wav, 1726,91s ˜1727,27s).

4.2 Mots et suites de mots fréquemment réduits

Dans la suite, nous présentons quelques mots et séquences de mots qui ont été fréquemment alignés à l'aide de suites de segments courts. Nous tentons d'identifier des phénomènes de réduction au-delà des segments. En effet, certains phénomènes de réduction peuvent se limiter à la simple chute d'un segment, comme la chute de la coda /l/ dans les mots « il », « ils », voire « elle », « elles ». Cependant, nous pensons que la réduction peut englober plusieurs segments (sous-jacents) consécutifs et résulter dans des formes de surface simplifiées, modifiées avec différents processus phonologiques à l'œuvre qu'il s'agit de mieux décrire dans le futur. Notre ambition ici se limite à essayer à mettre en évidence de tels phénomènes sur des exemples de mots et séquences de mots les plus représentatifs.

Nous présenterons tout d'abord les **mots** (c'est-à-dire sans considérer leur voisinage) qui ont été les plus fréquemment alignés avec des suites de segments courts (plus de 100 occurrences), avant de présenter des séquences de mots fréquemment alignées avec des suites de segments courts. Il faut noter qu'avec le critère appliqué (plus de 100 occurrences d'un même mot étiquetées « Alrt »), on ne peut tirer de conclusions que sur les mots fréquents.

Mot	**Alrt (occ.)**	**Total (occ.)**	$Taux\ (\%) = \frac{Alrt\ (occ.)}{Total\ (occ.)}$	**Détails** Seq. Occ.	Seq. Occ.	Seq. Occ.
voilà /vwala/	481	1723	28%	vwal 182 vwala 56	vwa 106 wala 24	wal 92 ala 21
crois /kʁwa/	129	486	27%	kʁwa 56	ʁwa 37	kʁw 36
trouve /tʁuv/	134	495	27%	tʁuv 84 uvə 2	ʁuv 32 tʁuvə 2	tʁu 18 ʁuvə 1

TABLE 2 – Mots alignés avec des suites de segments courts. « Alrt » et « Total » donnent respectivement le nombre d'occurrences du mot aligné avec des séquences de segments courts (Seq.) et le nombre total d'occurrences (Occ.). La colonne « Détails » précise les phones impliqués dans les « Alrt » alignés avec leur nombre d'occurrences.

Séquence de mots	Alrt (occ.)	Total (occ.)	$Taux\ (\%) = \frac{Alrt\ (occ.)}{Total\ (occ.)}$	Détails					
				Seq.	Occ.	Seq.	Occ.	Seq.	Occ.
par rapport /paʁ#ʁapɔʁ/	119	172	69%	aʁʁa	60	aʁʁ	46	paʁʁa	6
				ʁʁa	2	paʁʁ	2	ʁʁapɔʁ	1
par exemple /paʁ#ɛgzɑ̃pl/	117	239	49%	aʁɛgz	40	aʁɛg	24	ʁɛg	20
				ʁɛgz	8	aʁɛgzɑ̃p	6	paʁɛg	5
				aʁɛgzɑ̃	4	ʁɛgzɑ̃p	3	paʁɛgz	3
				aʁɛ	3	paʁɛ	1		
quand même /kɑ̃#mɛm/	177	765	23%	ɑ̃m	85	ɑ̃mɛm	63	kɑ̃m	18
				kɑ̃mɛm	4	kɑ̃mɛ	4	ɑ̃mɛmə	3

TABLE 3 – Séquences de mots alignées avec des suites de segments courts. « Alrt » et « Total » donnent respectivement le nombre d'occurrences de la séquence de mots alignée avec des suites de segments courts et le nombre total de ses occurrences. La colonne « Détails » donne les phones impliqués dans les « Alrt »avec leur nombre d'occurrences.

Le tableau 2 illustre des mots ayant plus de 100 occurrences alignés avec une suite d'au moins 3 segments courts (> 100 occurrences, au total 11 mots). Après vérification manuelle, nous avons pu constater que ces mots remplissent souvent un rôle de marqueur du discours. Ils ne sont pas très informatifs en soi, ils se comportent davantage comme une ponctuation (ex. voilà, je crois, je trouve, alors, tu vois, enfin, quoi) – que les locuteurs utilisent facilement dans un style de parole familier.

Le tableau 3 donne quelques séquences de mots fréquemment alignées avec une suite d'au moins 3 segments courts (> 100 occurrences, 6 séquences de mots). La même tendance a été observée que précédemment : il s'agit souvent de marqueurs discursifs et de « tics de langage ». Cependant, on y trouve également des expressions comme « par rapport » et « par exemple » qui ont les « taux d'alerte » les plus élevés parmi les séquences de mots trouvées. Il s'agit ici de deux expressions adverbiales qui restent parfaitement intelligibles même si elles sont prononcées d'une manière extrêmement réduite. On peut se demander comment sont réalisées de telles séquences de mots réduites. Quelles formes de surface sont produites ? Quels segments disparaissent ?

La figure 5 présente le spectrogramme avec la segmentation automatique de la locution « par exemple ». Une transcription phonétique manuelle donnerait à peu près [paʁɑ̃p]. Nous observons que le meilleur appariement entre signal et transcription a été obtenu grâce à la variante la plus courte sans le /l/ du cluster final /pl/, qui est la prononciation la plus adaptée selon le signal acoustique et le dictionnaire de prononciation. Ensuite, nous observons que la suite de phonèmes /ɛgzɑ̃/ a été alignée avec une suite de segments de durée minimale : il s'agit d'une réduction massive où la suite de segments [ɛgzɑ̃] a été prononcée comme un [ɑ̃].

La méthode ascendante mise en œuvre dans cette étude nous a permis de localiser des zones de parole à réduction massive. Les séquences de mots réduits ainsi mises en évidence révèlent qu'il s'agit avant tout de marqueurs discursifs. Mais on peut également y trouver des locutions polysyllabiques comme « par rapport » et « par exemple ». Ces locutions polysyllabiques se trouvent raccourcies en des réalisations de surface avec un nombre de syllabes plus faible. Dans des études futures, il serait intéressant d'analyser ces réductions en y intégrant une grille d'analyse prosodique dans la mesure où on peut avancer l'hypothèse que les syllabes non-accentuées seraient davantage sujettes au raccourcissement, voire à la disparition que les syllabes sous accent final.

5 Discussion

Cette étude sur la réduction temporelle montre que la combinaison de grands corpus avec l'alignement automatique présente des possibilités innovantes qui permettent de nouveaux points d'entrée d'analyses. Dans l'approche ascendante proposée, nous avons su profiter du « point faible » de l'alignement qui est la production d'une rafale de segments courts en zones temporellement réduites et nous avons pu apporter de nouvelles connaissances sur ce phénomène peu étudié auparavant. Dans la méthode ascendante appliquée dans cette étude, nous avons considéré comme « séquence réduite » les suites de segments d'au moins trois segments consécutifs de 30 ou 40 ms. Cette procédure très sélective nous a permis de localiser automatiquement les zones qui ont une forte probabilité d'être réduites (environ 14% du corpus). Nous avons constaté que les suites des segments courts impliquent souvent deux ou plusieurs mots (au lieu d'un seul mot). De plus, si la suite des segments courts fait partie d'un seul mot, elle est en général en position interne.

Nos résultats sur les mots et séquences des mots fréquemment réduits concernent surtout des locutions utilisées comme marqueurs discursifs qui ne portent que peu d'information sémantique et qui se comportent davantage comme une ponctuation ou qui peuvent éventuellement jouer un rôle dans la gestion des tours de parole lors de l'interaction.

En examinant les résultats sur la propension à la réduction des segments obtenus au niveau des mots, on peut remarquer que les séquences de segments réduits les plus fréquents partagent souvent des traits phonologiques (cf. « quand même » dans le tableau 3 pour lequel les séquences réduites les plus fréquentes sont toutes voisées, presque toutes nasales et les consonnes partagent le lieu d'articulation bilabiale) . Cette observation nous invite à étudier plus en détails le rôle des traits partagés dans une séquence de segments réduits sur la réduction en parole spontanée.

Il est intéressant cependant de noter que des locutions adverbiales polysyllabiques comme « par exemple » et « par rapport » apparaissent comme les plus réduites dans le corpus NCCFr. Ce résultat nous invite à rester vigilants sur la présence de réductions potentielles sur des mots ou locutions polysyllabiques moins fréquents. En effet, il faut garder à l'esprit qu'avec notre critère de sélection pour l'étude des mots et des séquences de mots (nous n'avons inclus que des séquences ayant plus de 100 occurrences étiquetées « Alrt »), on ne peut tirer de conclusions que sur les mots ou suites de mots les plus fréquents du corpus. Nous pensons que la réduction est certes favorisée par la fréquence des mots, mais qu'elle peut également être favorisée par d'autres régularités (ex. types de phonème impliqués, position syllabique, accentuation...) et des mécanismes (fusion, chute ou recombinaison de segments, de syllabes). Ce type de mécanismes peuvent potentiellement être à l'œuvre de manière similaire dans des phénomènes de réduction concernant des mots moins fréquents. Des études futures sur des plus grands corpus permettront de continuer ces pistes de recherche.

Remerciements

Ce travail est financé par Investissements d'Avenir – Projet Labex EFL (ANR-10-LABX-0083).

Références

Adda-Decker M. & Lamel L. (2000). The use of lexica in automatic speech recognition. In *Lexicon Development for Speech and Language Processing*, p. 235–266. Springer.

Adda-Decker M. & Lamel L. (2017). Discovering speech reductions across speaking styles and languages. In *Rethinking reduction : Interdisciplinary perspectives on conditions, mechanisms, and domains for phonetic variation*. Walter de Gruyter GmbH & Co KG.

Duez D. (1997). Acoustic markers of political power. *Journal of Psycholinguistic Research*, **26**(6), 641–654.

Duez D. (2003). Modelling aspects of reduction and assimilation of consonant sequences in spontaneous french speech. In *Proceedings of Spontaneous Speech Processing and Recognition, IEEE-ISCA*, p. 120–124 : University of Tokyo.

Ernestus M. T. C. (2000). *Voice assimilation and segment reduction in casual Dutch : A corpus-based study of the phonology-phonetics interface*. Thèse de doctorat, LOT, Utrecht.

Galliano S., Gravier G. & Chaubard L. (2009). The ester 2 evaluation campaign for the rich transcription of french radio broadcasts. In *Tenth Annual Conference of the International Speech Communication Association*.

Gauvain J.-L., Lamel L. & Adda G. (2002). The limsi broadcast news transcription system. *Speech communication*, **37**(1), 89–108.

Johnson K. (2004). Massive reduction in conversational american english. In *Spontaneous speech : Data and analysis. Proceedings of the 1st session of the 10th international symposium*, p. 29–54 : Tokyo, Japan : The National International Institute for Japanese Language.

Meunier C. & Espesser R. (2011). Vowel reduction in conversational speech in french : The role of lexical factors. *Journal of Phonetics*, **39**(3), 271–278.

Torreira F., Adda-Decker M. & Ernestus M. (2010). The nijmegen corpus of casual french. *Speech Communication*, **52**(3), 201–212.