



HAL
open science

Caractéristiques acoustiques et phonétiques des pauses remplies dans la conversation en français

Danielle Duez

► **To cite this version:**

Danielle Duez. Caractéristiques acoustiques et phonétiques des pauses remplies dans la conversation en français. Travaux interdisciplinaires du Laboratoire Parole et Langage, 2001, 20, pp.31-48. hal-00285404

HAL Id: hal-00285404

<https://hal.science/hal-00285404>

Submitted on 5 Jun 2008

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.

CARACTÉRISTIQUES ACOUSTIQUES ET PHONÉTIQUES DES PAUSES REMPLIES DANS LA CONVERSATION EN FRANÇAIS

Danielle Duez

Résumé

Dans cette étude, nous analysons certaines caractéristiques acoustiques et phonétiques des pauses remplies et leur relation au contexte prosodique dans la parole conversationnelle en français. Deux résultats majeurs émergent : 1) Absence de relation entre la durée des pauses remplies, le contour de la fréquence fondamentale (Fo), et les écarts entre les valeurs de Fo les plus hautes et les plus basses et 2) Absence de relation entre la valeur de Fo du début de la pause remplie et la valeur de Fo du sommet du syntagme prosodique, mais grande similarité entre la valeur de Fo du début de la pause remplie et de l'attaque non marquée du syntagme. Les valeurs de Fo du début de la pause remplie sont spécifiques au locuteur et liées à des contraintes physiologiques absolues, elles sont en revanche relativement stables au sein des extraits produits par un même locuteur. Ces caractéristiques suggèrent que les valeurs de l'attaque de la pause remplie sont utilisées par l'auditeur pour évaluer la gamme mélodique du locuteur.

Abstract

In the current analysis we examined the acoustic and phonetic characteristics of filled pauses in spontaneous French speech and their relationship to the prosody of the surrounding context. Two main results emerged : 1) There was no effect of the duration of filled pauses or their sentence location on their Fo patterns or on the differences between the highest and lowest values. 2) There was no relationship between peak-Fo values and the Fo values of filled-pause onsets, but the Fo values of filled-pause onsets and the Fo-values of non-marked breath-group onsets were highly similar. The Fo values of filled-pause onsets are speaker-dependent and strongly linked to the physiological, absolute aspects of speech production, but absolute within the same speaker's speech. It is assumed that filled-pause onset may be used by listeners as a reference for evaluating the speaker's pitch range.

Mots-clés : diffluences, pauses remplies, caractéristiques acoustiques et phonétiques, contexte prosodique, niveaux intonatifs.

Keywords : disfluencies, filled pauses, acoustic and phonetic characteristics, prosodic context, intonative levels.

DUEZ, D. (2001), Caractéristiques acoustiques et phonétiques des pauses remplies dans la conversation en français, *Travaux Interdisciplinaires du Laboratoire Parole et Langage*, vol. 20, p. 31-48.

I. Introduction

Ces vingt dernières années un intérêt grandissant s'est manifesté pour la production des diffluences telles que les pauses remplies (euh), les syllabes allongées (ne correspondant pas à l'allongement final de fin de syntagme), les répétitions (non sémantiques), les faux départs repris au non, les pauses silencieuses (le plus souvent distribuées à l'intérieur des syntagmes).

Les raisons en sont nombreuses : intérêt renouvelé pour la parole dite spontanée et toutes les manifestations de spontanéité (voir par exemple le numéro de la revue *Parole* consacré à la parole spontanée, 2001), investigation approfondie des processus d'encodage dans des situations de communication différentes, et nécessité d'améliorer les technologies de la parole, en particulier le naturel de la synthèse de parole et la fiabilité de la reconnaissance de parole.

Les pauses remplies sont probablement les diffluences qui ont soulevé la plus grande attention. Dès 1933, Bloomfield souligne leur existence et les classe dans les facteurs non linguistiques mais néanmoins pourvus de sens. Les pauses remplies sont les diffluences les plus courantes en français comme en anglais (Grosjean et Deschamps, 1972, 1973 et 1975) dans la parole des adultes (Grosjean et Deschamps, 1972, 1973 ; Duez, 1991) comme dans celle des adolescents (Candea, 2000 ; Sabin, Clemmer, O'Connell et Kowal, 1979). Les pauses remplies semblent liées aux difficultés de l'encodage, à la recherche lexicale et à la planification de l'énoncé.

De manière générale, les pauses remplies distribuées devant un mot lexical jouent un rôle dans les processus de sélection lexicale (Goldman-Eisler, 1968 ; Grosjean et Deschamps, 1972 et 1973 ; Maclay et Osgood, 1959), alors que les pauses remplies réalisées en frontière syntaxique ou au début d'un syntagme ont un rôle dans la programmation du syntagme à venir (Boomer, 1965 ; Duez, 1987 et 1991 ; Grosjean et Deschamps, 1972 et 1973 ; Maclay et Osgood, 1959). L'occurrence des pauses remplies dépend du degré d'incertitude du mot subséquent : elle est plus élevée devant les mots lexicaux que devant les mots grammaticaux (Maclay et Osgood, 1959), devant les mots abstraits que devant les mots concrets (Reynolds et Paivio, 1968). Il est à noter aussi que la fréquence des pauses remplies est plus grande à l'intérieur des syntagmes qu'à leur frontière, ce qui suggère qu'elles sont surtout dévolues à la fonction de recherche lexicale.

Les pauses remplies dépendent aussi du degré d'attention consciente portée au contenu du message (Christenfeld et Creager, 1996 ; Christenfeld, 1996). Elles sont nettement moins fréquentes dans des situations où le locuteur ne peut contrôler sa production (contrainte d'un rythme) et s'arrêter pour corriger les éventuelles erreurs dans la sélection lexicale. Ces résultats sont en accord avec le modèle de Levelt (1989) qui considère que les locuteurs contrôlent ce qu'ils disent à plusieurs niveaux. Les pauses remplies seraient donc produites

quand les locuteurs détectent une erreur ou une erreur à venir et s'arrêtent pour la corriger : elles sont des symptômes des processus de sélection de la parole.

Les caractéristiques acoustiques et phonétiques des pauses remplies et leur relation au contexte prosodique ont aussi fait l'objet d'études récentes. Leur objectif était d'examiner la manière dont les pauses remplies s'intègrent dans la structure intonative des phrases et/ou de définir un certain nombre d'indices fiables pouvant servir de critères pour la reconnaissance des pauses remplies. Dans ses prolégomènes aux structures intonatives, Léon (1969) note qu'en français les pauses remplies sont produites à une fréquence fondamentale spécifique au locuteur, et ce quelle que soit leur localisation dans l'énoncé. La valeur de la fréquence fondamentale des pauses remplies est similaire à la valeur moyenne de la fréquence fondamentale des syllabes inaccentuées et de l'attaque des unités prosodiques majeures. Dans son analyse acoustique des hésitations vocales, Guaitella (1991) met en évidence trois patrons prédominants pour les pauses remplies : (1) chute progressive, (2) chute irrégulière et (3) chute suivie d'une montée soudaine. O'Shaughnessy (1992) dans une étude portant sur la reconnaissance des hésitations en français ne relève pas de différences significatives entre les patrons des pauses remplies localisées à l'intérieur des syntagmes et à leur frontière : de manière générale, les pauses remplies ont un contour descendant ou plat. Il note cependant que les pauses remplies localisées aux frontières syntaxiques démarrent avec une valeur de F_0 plus élevée que les pauses remplies réalisées à l'intérieur des unités syntaxiques. Les pauses remplies marquées en frontières syntaxiques sont également plus longues que les pauses remplies situées à l'intérieur des syntagmes.

Shriberg et Lickley (1992) analysent les valeurs de la fréquence fondamentale de pauses remplies situées à l'intérieur des syntagmes et les valeurs de la fréquence fondamentale des pics précédant ces pauses dans des dialogues et des conversations en anglais britannique et anglais américain. Ils ne relèvent aucune dépendance entre la durée des pauses remplies et les variations de fréquence fondamentale. Ils montrent en revanche que les pauses qui ont les valeurs les plus élevées sont souvent associées aux pics les plus hauts. Pour ces auteurs, ces résultats sont la marque d'une relation systématique entre les pics et les pauses remplies.

Dans la présente étude nous nous proposons d'examiner certaines caractéristiques phonétiques et acoustiques des pauses remplies en relation avec la prosodie de leur contexte. Notre objectif est de vérifier si les pauses remplies sont produites à une valeur absolue spécifique au locuteur ou si elles sont dépendantes du contexte prosodique. Les pauses remplies sont examinées en fonction de leur distribution dans l'énoncé (à l'intérieur d'un mot prosodique ou à la frontière de groupes prosodiques majeurs et/ou mineurs) dans des

conversations produites par deux locuteurs et deux locutrices français. La valeur de Fo de l'attaque du groupe prosodique et du pic le plus proche de la pause remplie (précédent ou subséquent) constituent les mesures du contexte prosodique. Les mesures de la pause remplie sont celles de la fréquence fondamentale du début, du milieu (ou du point tournant) et de la fin de la pause. La durée de chaque pause remplie est également mesurée.

2. Procédure expérimentale

2.1. Sujets

Des conversations produites par deux locuteurs et deux locutrices d'origine et de langue maternelle françaises (une heure en tout) constituent la base de l'expérience. Les locuteurs et locutrices, âgés de 30 à 50 ans, sont de même niveau socioculturel. Ils (et elles) ont une audition et une production normales et aucune marque prononcée d'accent régional. Les locuteurs et locutrices, qui n'ont aucune idée du sujet de l'expérience, répondent de manière détendue et libre aux questions posées sur leur vie, leur enfance, leurs goûts et leurs projets. Les conversations sont enregistrées sur un magnétophone SONY à une vitesse de 19 cm/sec dans la chambre sourde du Laboratoire Parole et Langage d'Aix-en-Provence.

2.2. Transcription

Les conversations sont transcrites orthographiquement par l'auteur. La présence des pauses remplies et la réalisation des frontières sont contrôlées perceptivement par deux auditeurs.

2.3. Pauses remplies

Par pause remplie, nous entendons toute occurrence de "euh" [O]. Cette pause remplie peut être réalisée avant ou après une voyelle comme dans "euh on" [Oɔ̃], "mais euh" [m Eɔ̃] ou attachée à une consonne finale de mot suivie d'un schwa non prononcé comme dans "donc euh" [dɔ̃kɔ̃], "visite euh" [viziɔ̃]. En revanche, les cas où un schwa optionnel est réalisé et allongé comme dans "que" [kɔ̃:], "le" [lɔ̃:] et "de" [dɔ̃:] sont considérés comme des syllabes allongées et non des pauses remplies. Aucune distinction n'est faite entre les pauses remplies et les possibles "euh dits d'appui". Le "euh dit d'appui" dont l'existence est signalée par Fónagy (1989) est un phénomène en expansion qui paraît surtout avoir une fonction rythmique (Carton, 1999).

2.4. Localisation des pauses remplies

Les analyses prosodiques des discours (lus et spontanés) s'accordent à distinguer divers niveaux de frontières prosodiques. Par exemple, Silverman *et al.* (1992) proposent un système de notation comprenant quatre niveaux de frontières pour la transcription de l'intonation anglaise. Di Cristo (2000) distingue six unités différentes en parole spontanée (unité tonale, mot prosodique, unité intonative, syntagme intonatif, paraton mineur et paraton majeur), chacune d'entre elles est caractérisée par des contours prosodiques spécifiques et différents niveaux de frontières.

Une analyse des caractéristiques acoustiques et phonétiques des pauses remplies distribuées aux frontières des différentes unités prosodiques serait sans doute d'un grand intérêt pour la pragmatique. Nous avons cependant choisi d'en rester dans cette première étude des pauses remplies à deux localisations : intérieur d'un groupe mineur et frontière de groupe (majeur ou mineur). Le groupe mineur réfère ici au mot prosodique ou au groupe accentuel (Di Cristo, 1978-2000), il est constitué d'unités étroitement liées sémantiquement, syntaxiquement (par exemple un article et un nom, une préposition et un verbe). La pause remplie qui intervient à l'intérieur d'un groupe mineur est souvent réalisée devant un mot lexical. Les pauses remplies en frontières de groupe sont distribuées à la fin ou au début d'un groupe mineur ou d'un groupe majeur aussi nommé groupe intonatif (Di Cristo, 1978) ou groupe de souffle marqué (BG+) ou non (BG-) par une pause silencieuse (Vaissière, 1991). Il est à noter cependant que dans la parole spontanée, les frontières réalisées sont essentiellement celles des groupes majeurs. La présence de toute autre diffluence (correction, interruption, syllabe allongée) adjacente à la pause remplie est aussi contrôlée.

2.5. Niveaux et points de référence

Dans les études anciennes portant sur la prosodie du français parlé (voir Delattre, 1966b) quatre niveaux intonatifs sont habituellement utilisés pour représenter les contours intonatifs en français. Le niveau 2 est le niveau des attaques non marquées de syntagme, le niveau 3 est le niveau des syllabes finales de syntagmes à continuation mineure, le niveau 4 le niveau des syllabes finales de syntagmes à continuation majeure et le niveau 1 celui de la finalité. Un cinquième niveau peut être ajouté : c'est celui de l'emphase et des phrases exclamatives (Léon, 1971). Pour représenter les patrons intonatifs, Di Cristo (1976-78) distingue trois "points-clés" : l'attaque, la syllabe prétonique et la syllabe tonique. Les valeurs de la fréquence fondamentale de l'attaque peuvent être très élevées dans des phrases interrogatives, exclamatives, ou apportant une information nouvelle, cependant de manière générale, elles sont très stables

dans les unités déclaratives et très proches des valeurs de la fréquence fondamentale des syllabes prétoniques.

Il n'y a pas d'accent lexical en français (Togebly, 1965), mais une proéminence finale réalisée en frontière de groupe intonatif (Di Cristo, 1978) ou de groupe de souffle (Vaissière, 1991, 1995). Outre cette proéminence finale le français possède une proéminence initiale réalisée sur la première syllabe de mots lexicaux. Cette proéminence qui est attestée dès 1912 par Scherck (cité par Fónagy, 1980) paraît refléter un changement en cours (Fónagy, 1980) : elle dépend essentiellement du style (Fónagy et Fónagy, 1976 ; Lucci, 1983 ; Vaissière, 1974-75 ; Vihanta, 1993), du locuteur (Duez, 1978) et de la communauté sociale (Duez et Casanova, 1997). Contrairement à la proéminence finale qui est réalisée par un allongement, la proéminence initiale est marquée par une variation mélodique (Vaissière, 1974, 1975 ; Padeloup, 1990 ; Di Cristo, 1999).

Dans cette étude, les pics de fréquence fondamentale correspondent essentiellement à la réalisation d'une proéminence initiale ou finale. Ils peuvent également être la marque de l'accent dit emphatique qui se réalise très souvent par une forte montée de la F_0 (Garde, 1968 ; Duez, 1978 ; Di Cristo, 1999). Les valeurs de la fréquence fondamentale de l'attaque et du sommet sont essentielles dans la structuration prosodique de l'énoncé et dans l'intégration du registre du locuteur, elles sont donc prises comme valeurs de référence dans cette étude.

Les phrases du corpus sont numérisées à une fréquence d'échantillonnage de 16 kHz avec un ordinateur SUN. Les 170 pauses remplies du corpus sont sélectionnées à partir d'une écoute attentive des échantillons et les mesures sont effectuées à partir des spectrogrammes, des oscillogrammes et des courbes de fréquence fondamentale. La représentation phonétique des courbes de fréquence fondamentale est le résultat de l'application de l'algorithme MOMEL (Hirst et Essesper, 1993) qui permet de substituer à la courbe brute une courbe lisse et continue correspondant à la perception continue de la mélodie de la parole et qui fait apparaître les points-clés correspondant aux changements significatifs de la courbe. Les pauses remplies sont analysées acoustiquement et perceptivement. Chaque pause remplie est mesurée : ses limites sont la première et la dernière période de la voyelle [O]. Six valeurs de fréquence fondamentale sont enregistrées pour chacune des pauses remplies : F_0 du début, milieu et fin de la pause et F_0 de l'attaque et du sommet du syntagme, et de la syllabe suivant la pause remplie. Une illustration de l'étiquetage est présentée ci-dessous figure 1.

2.6. Mesures

Figure 1

De haut en bas on trouve l'oscillogramme, la fenêtre d'étiquetage, le spectrogramme et la courbe de mélodie modélisée (MOMEL) de la phrase " et il y avait des cours # euh de marxisme léninisme".

L'oscillogramme et le spectrogramme sont utilisés pour étiqueter les différents événements du signal de parole (début et maximum du syntagme, début, milieu et fin de la pause remplie). La courbe de Fo modélisée donne les valeurs (en Hz) des points-clés du syntagme et de la pause remplie. Dans l'exemple proposé, la valeur de Fo du pic est de 250 Hz.

2.7. Analyse

Les trois contours de base des pauses remplies sont les suivants : contour plat (P), montant (M) et descendant (D). La combinaison de ces trois patrons donne en tout neuf configurations (par exemple descendant (D), descendant+plat (DP), descendant+montant (DM). Un exemple de pause remplie ayant un contour descendant progressif peut être vu figure 2. Ces différents patrons sont examinés en fonction de la localisation de la pause remplie dans l'énoncé. Les valeurs initiales des pauses remplies sont comparées avec les valeurs de l'attaque du groupe dans lequel la pause remplie est réalisée et avec la valeur moyenne des attaques de groupes non marquées. La valeur de la fréquence fondamentale du pic précédent ou subséquent est mesurée pour chacune des pauses remplies. Les valeurs initiales des pauses remplies sont comparées avec les valeurs obtenues pour l'ensemble des pics.

Figure 2

*Exemple de pause remplie à contour de fréquence fondamentale descendant progressif (le plus fréquent).
En ordonnée, on trouve les valeurs de fréquence fondamentale (en Hz),
en abscisse, la durée (en dixième de ms).*

3. Résultats

3.1. Patrons de fréquence fondamentale des pauses remplies

	D	DP	DM	P	PD	PM	M	MD	MP	T
I	35	14	9	10	1	4	2	5	1	81
F	29	22	10	13	4	2	3	5	1	89
T	64	36	19	23	5	6	5	10	2	170

Tableau 1

Nombre de pauses remplies en fonction de leur localisation dans l'énoncé

(I: intérieur d'un mot prosodique; F: en frontière de deux mots prosodiques ou de deux groupes mineurs ou majeurs) et patron de la pause remplie and Fo pattern (D: descendant, DP: descendant+plat, DM: descendant-montant, P: plat, PD: plat+descendant, PM: plat+montant, M: montant, MD: montant+descendant et MP: montant+plat). T signifie total.

Les résultats qui sont reportés au tableau 1 montrent une répartition similaire des différents patrons des pauses remplies à l'intérieur des syntagmes et en frontière de mot prosodique et groupes majeurs et mineurs. Le test du χ^2 ne révèle aucune interaction significative entre la distribution des pauses remplies et leur configuration ($\chi^2: 0.7, df(8), n.s.$).

Contours des pauses remplies. Les configurations descendantes de fréquence fondamentale sont les plus fréquentes, soit 105 sur 170 en tout. La configuration pour ces 105 pauses remplies est la suivante : 64 d'entre elles ont une descente continue et progressive, 36 présentent une descente suivie d'un plateau, 5 un plateau suivi d'une descente. Pour 23 autres pauses remplies, il y a un contour totalement plat enfin, pour les 37 pauses remplies restantes il y a combinaison d'une montée et d'un plateau, d'une montée et d'une descente (ou réciproquement).

Calcul de la pente de F_0 . De manière générale, les différences de fréquence fondamentale entre la valeur initiale et la valeur finale des pauses remplies ont un pourcentage inférieur à 20%, ce qui est en complet accord avec les pourcentages obtenus par Guaitella (1991) et O'Shaughnessy (1992). Cependant, 36 des 170 pauses remplies ont des différences entre la valeur initiale et la valeur finale plus élevées (allant de 25% à 50%). La majorité de ces pauses exhibe des chutes brutales et abruptes (16 sur 36) ou des patrons descendant+plat (8 sur 36). Pour les 12 pauses remplies restantes la répartition est la suivante : descendant+montant (2), plat+descendant (1), plat+montant (2), montant (3), montant+descendant (4). L'interaction

entre les patrons de fréquence fondamentale et les différences de fréquence fondamentale est significative ($\chi^2 = 19.8$, $p = 0.01$).

Les pauses remplies qui présentent des différences entre les valeurs initiale et finale supérieures à 25% sont localisées tout aussi bien à l'intérieur de mots prosodiques (22) qu'aux frontières de groupes mineurs ou majeurs (3 au début et 11 à la fin). Il y a en outre présence d'une pause silencieuse pour 14 de ces 36 pauses remplies. Certaines de ces montées abruptes ou chutes brutales (elles peuvent atteindre des écarts proches de 60 Hz) sont dues à la voix craquée (neuf cas pour les hommes et huit cas pour les femmes).

3.2. Durée des pauses remplies

	D	DP	DM	P	PD	PM	M	MD	MP	T
I	432	669	468	468	449	428	500	666	202	476
B	450	501	577	478	376	491	485	313	1000	519
T	441	585	522	473	412	459	492	489	601	

Tableau 2

Durée des pauses remplies (en ms) en fonction de leur localisation dans l'énoncé (I : intérieur d'un mot prosodique ; F : en frontière de deux mots prosodiques ou de deux groupes majeurs ou mineurs) et patron de la pause remplie and Fo pattern (D : Descendant, DP : descendant+plat, DM : descendant+montant, P : plat, PD : plat+descendant, PM : plat+montant, M : montant, MD : montant+descendant et MP : montant+plat). T signifie total.

Une analyse de variance (ANOVA) réalisée sur la durée des pauses remplies ne révèle aucun effet significatif du patron de fréquence fondamentale [$F(152,8) = 1.2$, $p = 0.2$], et de la localisation dans l'énoncé [$F(152,1) = 0.5$, $p = 0.4$], mais un effet significatif de l'interaction de ces deux facteurs [$F(152,8) = 2.1$, $p = 0.03$].

3.3. Effet de la durée sur les différences de F_0 entre la valeur la plus basse et la valeur la plus élevée des pauses remplies

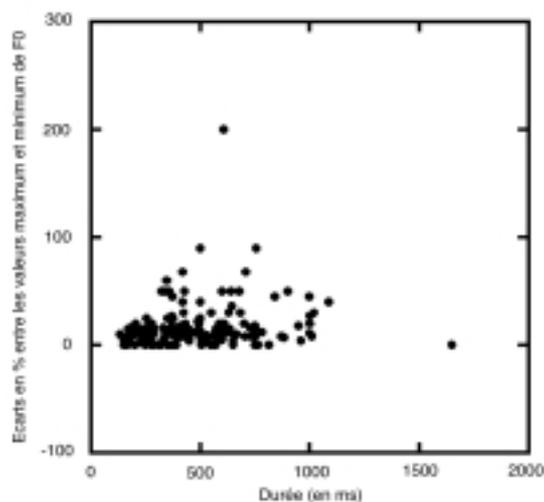


Figure 3
Différences de F_0 (exprimées en %) entre les valeurs extrêmes de la pause remplie en fonction de la durée

Les résultats reportés dans la figure 3 ne révèlent aucun effet de la durée sur les différences (exprimées en %) entre la valeur la plus élevée (le plus souvent la valeur initiale) et la valeur la plus basse (le plus souvent la valeur de la fin de la pause remplie). La valeur du coefficient de corrélation r^2 (0.018) confirme l'absence de relation entre la durée et les différences de fréquence fondamentale ainsi que la totale indépendance de la durée et des patrons de fréquence fondamentale des pauses remplies. Cette indépendance est une caractéristique des pauses remplies qui viennent s'ajouter au contour en fin de syntagme : elle les distingue des contours intonatifs où les variations de fréquence fondamentale sont étroitement corrélées à la durée (Lyberg, 1977).

3.4. Valeurs initiales (en Hz) des pauses remplies et des attaques des syntagmes

	Hommes		Femmes	
	Loc1	Loc2	Loc3	Loc4
Attaque des PR	110	114	170	176
	(15)	(14)	(32)	(14)
Attaque de G	116	116	190	186
	(12)	(20)	(39)	(20)
Attaque des G non marquées	116	115	179	179
	(15)	(18)	(30)	(11)

Tableau 3

Valeurs moyennes de F_0 de l'attaque des pauses remplies (PR) et attaque des groupes (G). Deux valeurs différentes sont reportées pour les attaques de groupes de souffle : (1) celles dont des groupes dans lesquels les pauses remplies sont localisées et (2) celles des groupes majeurs non marquées (sont exclues les valeurs des phrases exclamatives, interrogatives et à contenu informationnel nouveau). Les valeurs des écarts-types sont notées entre parenthèses.

Une analyse de variance (ANOVA) révèle des différences significatives entre les valeurs de fréquence fondamentale des pauses remplies et des attaques de groupes dans lesquels les pauses sont localisées [$F(1, 306)=11, p=0.001$], un effet significatif du locuteur [$F(3,306)=235, p=0.0001$], et une interaction significative entre les deux facteurs [$F(3,306)=3.2, p=0.02$]. Les différences sont probablement dues au fait que les attaques de certains groupes sont sémantiquement ou syntaxiquement marquées et exhibent des valeurs de fréquence fondamentale élevées (par exemple, les plus hautes valeurs de Loc2 et Loc3 sont respectivement de 220 Hz et 300 Hz). En revanche, la comparaison des valeurs de fréquence fondamentale des pauses remplies et des attaques de groupes non marquées ne révèle aucune différence significative [$F(1,343)=3, p=0.06$], ce qui confirme l'hypothèse faite plus haut. Les valeurs moyennes de fréquence fondamentale des pauses remplies et des attaques de groupes marquées et non marquées sont reportées au tableau 3.

3.5. Attaques des pauses remplies et pics des syntagmes

Les résultats présentés figure 4 confirment l'absence de corrélation qui existe entre les valeurs initiales des pauses remplies et les valeurs des pics ($r^2=0.04$). Les valeurs initiales des pauses remplies sont groupées autour de la moyenne et sont très stables. On peut observer cependant quelques exceptions avec des valeurs basses (60 Hz, 80 Hz, and 60 Hz pour Loc1, Loc2 and

Loc3, respectivement) dues à des effets physiologiques (voix craquée). Ces valeurs basses caractérisent essentiellement des pauses remplies qui débutent avec une montée abrupte et des pauses remplies localisées entre deux pauses silencieuses. La figure 5 illustre cette absence de corrélation.

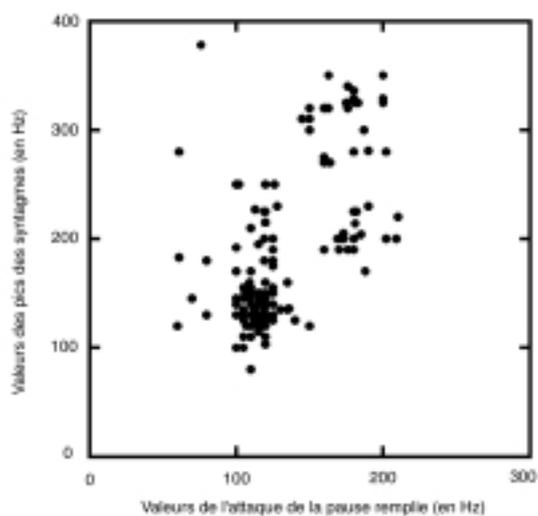


Figure 4

Valeurs de F_0 de l'attaque de la pause remplie en fonction des valeurs des pics des syntagmes

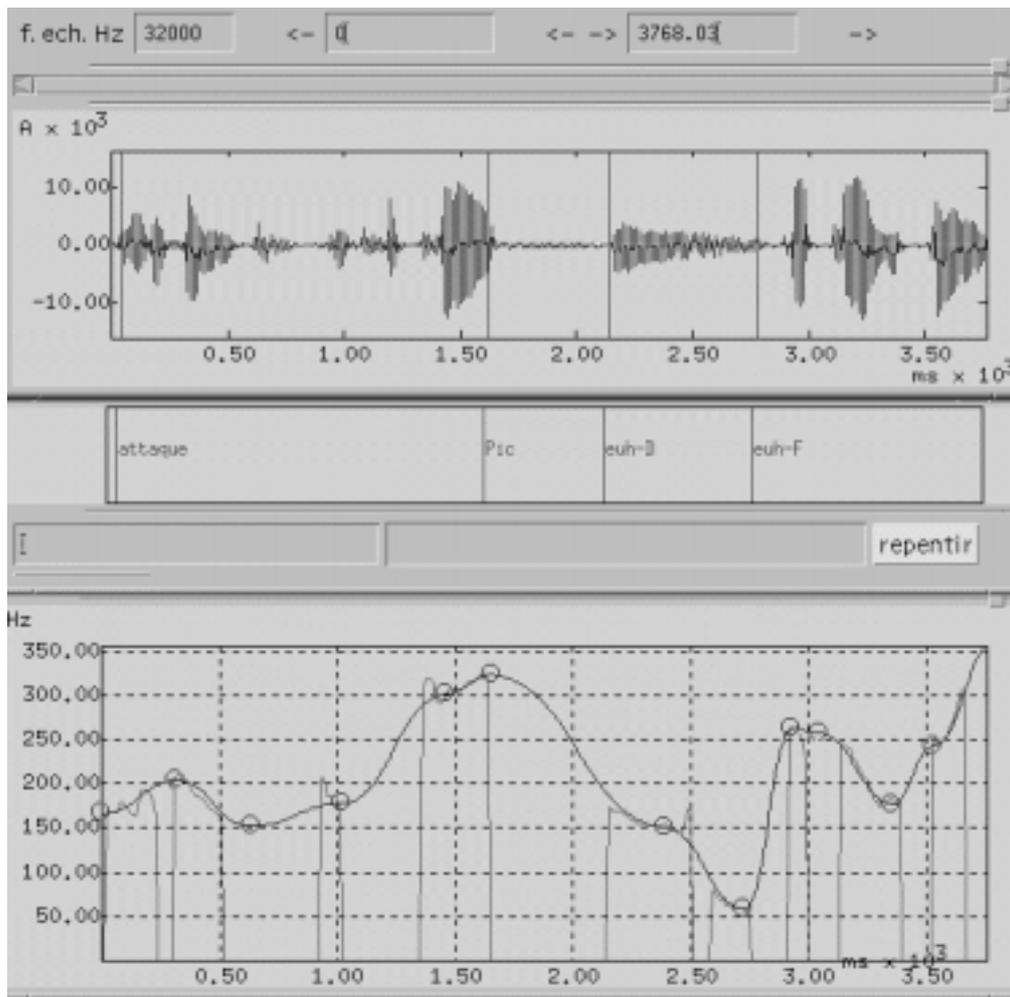


Figure 5

Cet exemple est une illustration de l'absence de corrélation entre les valeurs initiales des pauses remplies et celles des pics. Dans la phrase " donc c'est très # euh coûteux ", l'attaque du syntagme est à 170 Hz environ, le pic à 200 Hz environ, l'attaque de la pause remplie qui est localisée à l'intérieur du syntagme et suit immédiatement un silence est également à 170 Hz environ.

4. Conclusions

La présente analyse permet de tirer les conclusions suivantes :

- 1) Il n'y a aucun effet de la durée des pauses remplies et de leur localisation dans le syntagme

sur leur configuration générale et sur leurs valeurs de fréquence fondamentale les plus hautes et les plus basses.

2) Il n'y a aucune relation entre les valeurs de fréquence fondamentale des pics du syntagme et de l'attaque des pauses remplies, il existe en revanche une grande similarité entre la valeur initiale de la pause remplie et la valeur de l'attaque non marquée du syntagme.

Comme nous l'avons déjà souligné plus haut, le niveau 2 est un point-clé fondamental pour la perception et la production des unités intonatives (Di Cristo, 1976 et 1978). C'est une valeur de référence pour l'évaluation de la gamme mélodique ("pitch range") du locuteur et par conséquent une ligne de base pour l'intégration des différents points-clés. Les pauses remplies ont leur point d'ancrage au niveau 2 : elles peuvent donc constituer une référence pour l'intégration des différents niveaux. Cette interprétation est en accord avec la proposition de Léon (1969) : prendre les valeurs de fréquence fondamentale des pauses remplies pour définir les valeurs de référence du niveau 2. Réciproquement, les valeurs de fréquence fondamentale des attaques non marquées de syntagmes peuvent être utilisées comme critères dans la reconnaissance des pauses remplies.

L'ancrage des pauses remplies au niveau de l'attaque contredit l'idée commune selon laquelle les pauses remplies seraient des temps de repos. On sait que les attaques vocales mettent en jeu l'activité respiratoire et l'activité des muscles de la glotte de manière à produire les premières vibrations d'un son (Scotto di Carlo, 1975). Quatre phases peuvent être définies au niveau physiologique : tension des muscles respiratoires et des muscles du larynx permettant au larynx d'aller de sa position de repos à sa position phonatoire, 2) début de l'expiration, 3) fin de l'adduction des cordes vocales et 4) initiation de la vibration des cordes vocales (notons d'ailleurs que l'absence d'ajustement entre les phases 2 et 3 résultent souvent en voix craquée, qui est une caractéristique fréquente des pauses remplies). Il paraît donc y avoir une étroite correspondance entre l'activité physiologique et l'activité cognitive dans la production des pauses remplies, entre le positionnement des organes de la parole et la programmation des unités lexicales et syntaxiques.

De manière générale, les valeurs initiales de F_0 des pauses remplies sont similaires à celles de l'attaque des syntagmes au sein des extraits produits par un même locuteur. Elles varient en revanche d'un locuteur à l'autre : elles sont donc liées à des contraintes physiologiques propres au locuteur. L'absence de corrélation entre les valeurs initiales de F_0 des pauses et les valeurs

de Fo des pics observée ici est en opposition avec l'étroite corrélation observée pour l'anglais britannique et l'anglais américain (Shriberg et Lickley, 1993). Les divergences observées peuvent être dues à des caractéristiques spécifiques au français et à l'anglais : de manière générale, les différences sont réalisées dans des contextes phonologiques, prosodiques, syntaxiques et sémantiques spécifiques aux langues (Fox Tree, 1995 ; Grosjean et Deschamps, 1975).

Il est à noter qu'une minorité de pauses remplies exhibe des valeurs plus hautes que les valeurs des attaques non marquées des groupes. Ces pauses peuvent avoir une fonction pragmatique et être utilisées à des fins communicatives. En ce sens, elles pourraient être une manifestation de l'un des aspects relatifs et fonctionnels de la parole (Crystal, 1971). Une analyse des pauses remplies dans un corpus plus vaste d'échantillons de différents styles de parole spontanée, produits par un nombre plus élevé de locutrices et locuteurs, en relation avec les différentes unités prosodiques de la parole spontanée (Di Cristo, 2000) paraît nécessaire. Elle devrait permettre de valider les résultats préliminaires reportés dans cet article et d'obtenir une meilleure compréhension du rôle des pauses remplies dans la communication.

Bibliographie

BLOOMFIELD, L., (1933) *Language*, New York.

BOOMER, D.S. (1965) Hesitation and grammatical encoding, *Language and Speech*, 8, p. 148-158.

CANDEA, M. (2000) *Contribution à l'étude des pauses silencieuses et des phénomènes dits "d'hésitation" en français oral spontané*, thèse de doctorat nouveau régime, Paris III, manuscrit non publié.

CARTON, F., (1999) L'épithèse en français contemporain: étude phonétique, *Faits de Langues: Oral-Ecrit: Forme et Théorie*.

CHRISTENFELD, N.J.S. (1996) Effects of a metronome on the filled pauses of fluent speakers, *Journal of Speech and Hearing Research*, 39, p. 1232-1238.

CHRISTENFELD, N.J.S. & CREAGER, B. (1996) Anxiety, alcohol, aphasia and ums. *Journal of Personality and Social Psychology*, 70, p. 451-460.

CRYSTAL, D. (1971) Relative and absolute in intonation analysis. *Journal of the International Phonetic Association*, 1 (1), p. 17-28.

DELATTRE, P. (1966a) La leçon d'intonation de Simone de Beauvoir. Etude d'intonation déclarative comparée, *Studies in French and comparative phonetics*, Mouton & CO : The Hague, p. 75-83.

DELATTRE, P., (1966b) Les dix intonations de base du français, *The French Review*, 40(1), p. 1-14.

- DI CRISTO, A. (1976) Des traits acoustiques aux indices perceptuels. Application d'un modèle d'analyse prosodique à l'étude du vocatif en français, *Travaux de l'Institut de Phonétique d'Aix-en-Provence*, 3, 1976, p. 213-358.
- DI CRISTO, A., (1978) *De la microprosodie à l'intonosyntaxe*, Thèse de doctorat d'état, Université de Provence, 1978.
- DI CRISTO, A. (1999) Le cadre accentuel du français contemporain : Essai de modélisation. *Langues*, 2 (3), p. 184-205 et *Langues*, 2 (4), p. 258-267.
- DI CRISTO, A. (2000) La problématique de la prosodie dans l'étude de la parole dite spontanée, *Parole*, 15/16, p. 189-251.
- DI CRISTO, A. (à paraître) Vers une modélisation de l'accent en français, *Journal of French Language Studies*.
- DUEZ, D. (1978) *Essai sur la prosodie du discours politique*, Thèse de Doctorat de 3ème cycle, Paris III.
- DUEZ, D. (1987) *Contribution à l'étude de la structuration temporelle de la parole en français*, Thèse de Doctorat d'état, Université de Provence, 1987.
- DUEZ, D. (1991) *La pause dans la parole de l'homme politique*, Editions du CNRS, collection Sons et Parole, Paris.
- DUEZ, éd. invitée, (2000 et 2001) Parole spontanée, *Parole*, vol. 15, 16, 17, 18, 19.
- DUEZ, D. & CASANOVA-ROSSI, M.H., (1998) Quelques aspects de l'organisation temporelle du parler des banlieues parisiennes, *Parole*, 1, p. 59-73.
- FONAGY, I. (1980) L'accent français : accent probabilitaire, In I. Fónagy and P. Léon (eds), *L'accent en Français contemporain*, *Studia Phonetica* 8, p. 53-97.
- FONAGY, I. (1989) Le français change de visage ? *Revue Romane*, 24 2, p. 225-254.
- FONAGY, I & FONAGY, J. (1976) Prosodie professionnelle et changements prosodiques, *le Français Moderne* 44, p. 414-420.
- FOX TREE, J. E., (1995) The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech, *Journal of Memory and Language*, 34, p. 709-738.
- GARDE, P. (1968) *L'accent*, PUF : Paris.
- GOLDMAN-EISLER, F. (1968) *Experiments in spontaneous speech*, The Academic Press.
- GROSJEAN, F. & DESCHAMPS, A. (1972) Analyse des variables temporelles du français spontané, *Phonetica*, 26, p. 129-149.
- GROSJEAN, F. & DESCHAMPS, A. (1973) Analyse des variables temporelles du français spontané, *Phonetica*, 28, p. 191-226.
- GROSJEAN, F. & DESCHAMPS, A., (1975) Analyse contrastive des variables temporelles de l'anglais et du français: vitesse de parole et variables composantes :phénomènes d'hésitation, *Phonetica*, 31, p. 144-84.
- GUAÏTELLA, I. (1991) *Rythme et parole : comparaison critique du rythme de la lecture oralisée et de la parole spontanée*, thèse de doctorat, Université de Provence.
- HIRST, D.J. & ESPESSER, R. (1993). Automatic modelling of fundamental frequency using a quadratic spline function. *Travaux de l'Institut de Phonétique d'Aix*, 15, p. 71-85.
- LEON, P. (1969) Prolégomènes à l'étude des structures intonatives, *Studia Phonetica* 2 (P. Léon and P. Martin, eds) Montréal, Paris and Bruxelles : Didier.

- LEON, P. (1971) *Essais de phonostylistique*, Didier : Montréal, Bruxelles.
- LEVELT, W.J.M., (1989) *Speaking : From Intention to Articulation*, MA: MIT Press.
- LUCCI, V. (1983) *La variation situationnelle en français standard*. Grenoble : Presses de l'Université.
- LYBERG, B. (1977) Final lengthening –partly a consequence of restrictions on the speech of fundamental-frequency changes ? *Journal of Phonetics*, 7, p. 187-196.
- MACLAY, H. & OSGOOD, C.E. (1959) Hesitation Phenomena in Spontaneous English Speech, *Word*, 1, p. 19-43.
- MAHL, G. F.(1956) Disturbances in the patient's speech in psychotherapy, *Journal of Abnormal and Social Psychology*, 42, p. 3-32.
- O'SHAUGHNESSY, D. (1992) Recognition of hesitations in spontaneous speech, *IEEE*, 1, p. 521-524.
- PASDELOUP, V. (1990) *Modèles de règles rythmiques du français appliqué à la synthèse de parole*. Thèse de doctorat, Université de Provence.
- REYNOLDS, A.& PAIVIO, A., (1968) Cognitive and emotional determinants of speech, *Canadian Journal of Psychology*, 22, p. 164-175
- SABIN, E.J., CLEMMER, E.J., O'CONNELL, D.C. KOWAL, S., A (1979) Pausological approach to speech development, in A. Siegman and S. Feldstein (eds), *Of speech and time : temporal speech patterns in interpersonal context*, New Jersey : Hillsdale, p. 35-55.
- SCHERCK, O. (1912) *Über den Französischen Akzent*, Dissertation Berlin, Schmsow, Kirchstrasse, n°1.
- SCOTTO DI CARLO, N. (1975) Le mécanisme des attaques dans le chant, *Cahiers de linguistique, d'orientalisme et de slavistique* (Mélanges à G. Mounin), 5-6, p. 365-371.
- SILVERMAN, K., BECKMAN, M.E., PITRELLI, J., OSTENDORF, M., WIGHTMAN, C., PRICE, P., PIERREHUMBERT, J. & HIRSCHBERG, J. (1992). ToBI : a standard for labelling English prosody. *Proceedings of the Second International Conference on Spoken Language Processing*, 2, p. 867-870.
- SHRIBERG, E. & LICKLEY, R. (1993) Intonation of clause-internal filled pauses, *Phonetica*, 50, p. 172-179.
- TOGEBY, K. (1965) *Structure immanente de la langue française*, Paris, Larousse.
- VAISSIÈRE, J. (1974) On French Prosody, *Quarterly Progress Report*, Massachusetts Institute of Technology, Research laboratory of Electrics, 114, p. 212-223.
- VAISSIÈRE, J. (1975) Further note on French Prosody, *Quarterly Progress Report*, Massachusetts Institute of Technology, Research laboratory of Electrics, 115, p. 251-261.
- VAISSIÈRE, J. (1991) Rhythm, accentuation and final lengthening in French, in Sundberg, J., Nord, L. and Carlson, R. (eds), *Music, Language and Brain*, Macmillan, Houndsmills, p. 108-120.
- VAISSIÈRE, J. (1995) Phonetic explanations for cross-linguistic prosodic similarities, *Phonetica*, 52, p. 123-130.
- VIHANTA, (1993) Focalisations et autres proéminences en français lu et spontané, *Mélanges offerts à L. Lindgren*, Turun Ylopisto, p. 258-289.