

Pauses et hésitations en français spontané

Estelle Campione & Jean Véronis

Equipe DELIC

Université de Provence, 29 av. R. Schuman – 13621 Aix-en-Provence Cedex 1, France

Tél.: ++33 (0)4 42 95 31 37 - Fax: ++33 (0)4 42 95 34 95

Mél: Estelle.Campione@up.univ-aix.fr - <http://www.up.univ-mrs.fr/delic>

ABSTRACT

In traditional terminology, *silent* and *filled pauses* are grouped together, whereas *hesitation lengthening* is put into a separate category. However, while these various phenomena are very often linked, there have been few studies on how they interact. We analyzed an hour of spontaneous speech to show that silent and filled pauses operate in a totally different way, and that contrary to common belief, silent pauses by themselves never serve as hesitation marks, but only do so when coupled with other marks - mostly syllabic lengthening and filled pauses. These last two hesitation marks have similar acoustic and articulatory characteristics, and are distributed and function alike.

1. INTRODUCTION

On distingue de façon traditionnelle deux sortes de pauses (cf. par exemple Duez [7]) : les pauses *silencieuses*, dans lesquelles toute production vocale s'interrompt, à l'exception éventuellement de bruits respiratoires, et les pauses *remplies* (ou *sonores*), constituées d'un item quasi-lexical (*eah* en français, *er/erm* en anglais).

Regrouper ces phénomènes acoustiquement et articulatoirement très différents sous le même terme de *pause* revient à poser une hypothèse implicite quant à une identité de fonction. Or, rien n'est moins sûr qu'une telle identité. On a noté depuis bien longtemps (par exemple Boomer [2]) que les pauses silencieuses ont un double rôle. Certaines pauses silencieuses sont *démarcatives*, et apparaissent à la jonction de segments du discours, à la structuration desquels elles participent. D'autres sont des pauses *d'hésitation*, et sont le reflet de difficultés que rencontre ponctuellement le locuteur dans ses opérations mentales de « recherche et d'encodage » (Barik [1]) ou du « travail de formulation » (Morel & Danon-Boileau [14]) liées à la production du discours. A l'inverse, il semble que les pauses remplies soient cantonnées exclusivement à ce deuxième rôle : elles constituent en quelque sorte un signal conventionnel de la part du locuteur, lui permettant d'occuper le terrain de l'interaction et d'éviter d'être interrompu pendant le laps de temps nécessaire à la construction de la suite de son énoncé (cf. Clark & Clark [5], etc.).

Il est intéressant de remarquer que le rôle de signal d'hésitation correspondant aux pauses remplies est également dévolu à certains allongements syllabiques (qui affectent généralement une voyelle en fin de mot), que nous appellerons *allongements d'hésitation*. On a déjà noté que leurs propriétés sont proches de celles des pauses remplies (cf. Guaïtella [11], Candéa [4]), au point que les mêmes algorithmes permettent de détecter les deux (cf. Goto, Itou, & Hayamizu [10]). La terminologie établie regroupe donc des phénomènes acoustiquement et fonctionnellement très différents sous le même terme de *pause*, alors qu'elle met dans des catégories séparées les allongements d'hésitation et les *eah* malgré leur similitude de propriétés et de fonctionnement. Nous regrouperons ci-après, comme le font Goto et al. [10], allongements d'hésitation et *eah* sous le même terme de *pauses remplies*.

Il existe dans la littérature un grand nombre d'études sur les pauses silencieuses (voir par exemple l'état de l'art de Zellner [17]). Par contre, comme le montre une thèse récente (Candéa [4]), un nombre beaucoup plus restreint de travaux se sont intéressés aux marques du travail de formulation en général, aux pauses remplies et aux allongements en particulier (sans doute ceci est-il dû au fait, comme le font remarquer Cutler [6], Duez [8], etc., que la majorité des travaux phonétiques ont été consacrés à la « parole de laboratoire », au détriment de l'oral spontané). Il n'existe en tout cas presque pas de travaux sur l'interaction entre ces différents phénomènes. Or, comme nous le montrons ci-après, ils sont très souvent associés.

Nous proposons dans cette communication de fournir une étude précise des interactions entre pauses silencieuses, pauses remplies et allongements d'hésitation, à partir d'un échantillon du *Corpus de référence de français parlé* composé de 8 500 mots et 54 minutes de parole, faisant intervenir 10 locuteurs différents (5 hommes et 5 femmes). Nous montrons que les pauses silencieuses, hormis les pauses très brèves, n'ont jamais à elles seules un rôle de marque d'hésitation, et que pour assurer une telle fonction, elles sont obligatoirement associées à une autre marque du travail de formulation, typiquement un allongement ou un *eah* (ou une combinaison de ces éléments).

2. PAUSES SILENCIEUSES

2.1. Annotation

Les pauses silencieuses ont été détectées au moyen d'un programme qui calcule la fréquence fondamentale¹, et repère l'absence de voisement pendant une période donnée. Nous avons appliqué un seuil de 200 ms, en accord avec les études ultérieures (Candéa [4]), sachant que les pauses plus courtes, dont l'existence et l'importance ont été soulignées dans la littérature (Hieke, Kowal & O'Connell [12]), seraient ajoutées manuellement. Nous avons ensuite corrigé la totalité des pauses silencieuses à l'aide de l'éditeur de signal MES : les pauses détectées à tort (correspondant généralement à des occlusives sourdes) ont été supprimées, celles qui n'ont pas été détectées ont été ajoutées (y compris en dessous du seuil initial de 200 ms), et les frontières de celles qui ont été bien repérées ont été ajustées si besoin. Le corpus comportait 1375 candidats-pauses détectés et 1163 pauses effectives après correction. Nous avons catégorisé les pauses silencieuses en pauses brèves (< 200 ms), moyennes (200-1000 ms) et longues (> 1000 ms), notées respectivement ^, + et ++.

2.3. Rôle démarcatif

Il semble largement accepté par les psycholinguistes que la production de la parole relève d'un cycle planification-exécution, qui résulte en une succession d'unités relativement brèves (que nous appellerons ci-après *segments discursifs*), séparées par des pauses silencieuses (Fromkin [9] ; Levelt [13] ; etc.). Les pauses silencieuses sont nécessaires à la fois pour la planification des énoncés par les locuteurs, et le traitement par les auditeurs. Au contraire des pauses d'hésitation, ces pauses démarcatives ont donc un rôle important dans la structuration du discours.

Nous avons marqué dans le corpus toutes les pauses silencieuses de nature démarcative. Ces pauses sont assez faciles à repérer car de multiples indices convergent (intonation, qualité et durée vocalique, syntaxe, etc.). Nous avons en particulier utilisé la présence d'un mouvement mélodique majeur, ascendant ou descendant, mesuré après détection automatique et correction manuelle (Campione [3]). L'exemple ci-dessous montre le type de découpage obtenu (les segments discursifs sont séparés par || et l'étoile marque un accent de focalisation) :

ben je travaille dans un pressing ↗ ++ || on fait pas que le pressing on fait aussi la blanchisserie ↗ + || plus la blanchisserie d'ailleurs ↘ + || les draps les nappes la restauration ↗ ++*

Le Tableau 1 donne la répartition des différents types de pauses silencieuses. On constate que les pauses silencieuses sont démarcatives dans 71% des cas.

Tableau 1 : Types de pauses silencieuses

Longueur	Démarcative	Non-démarcatives	Total
brèves	0	24	24
moyennes	673 (70%)	293 (30%)	966
longues	148 (86%)	25 (14%)	173
Total	821 (71%)	342 (29%)	1163

Les pauses brèves ne sont jamais démarcatives, ce qui semble justifier les seuils de l'ordre de 200 ms utilisés dans certaines études. Nous rejoignons la position de Morel & Danon-Boileau [14] pour lesquels les pauses en dessous de ce seuil n'auraient pas de valeur iconique définie. Elles sont souvent de nature respiratoire et interviennent généralement en des points où la cohésion intonative est patente :

donc c'est ^ la calandre qui travaille aussi

3. PAUSES REMPLIES

3.1. Annotation

Allongements d'hésitation et *euh* (que nous regroupons sous le terme de pauses remplies) se caractérisent par une voyelle continue de durée très supérieure à la normale, de qualité vocalique constante et sont associées à une fréquence fondamentale (F_0) plate ou très légèrement descendante (Guaïtella [11]). Ces caractéristiques semblent être communes à de nombreuses langues (Quimbo, Kawahara & Doshita [15] ; Goto et al. [10]).

A part la longueur de la tâche, le marquage des pauses remplies dans les corpus présente donc assez peu de difficultés, sous réserve d'une écoute très soignée. Les allongements d'hésitation se distinguent clairement des allongements liés à la structure syntaxique, aux accents, etc., qui ont généralement, du moins en français, un contour intonatif ascendant ou descendant (cf. Vaissière [16]). De plus, les allongements d'hésitation portent très souvent sur des mots-outils, ou se situent à des positions qui ne constituent pas des points de rupture syntaxique ou discursive. Les *euh* sont des items quasi-lexicaux facilement reconnaissables, et il n'y a pratiquement jamais désaccord entre annotateurs sur leur présence. Seuls quelques cas sont difficiles à trancher sur des mots terminant par un schwa, et pour lesquels on peut hésiter entre un allongement du schwa ou un passage progressif à un *euh*. La question semble totalement indécidable, aussi bien à l'écoute qu'au vu des courbes mélodiques et sonagrammes, et on peut se demander si cette « hésitation » sur l'hésitation n'est pas en fait révélatrice de l'identité de fonction de l'allongement et du *euh*. Nous notons les hésitations sous forme d'allongement syllabique au moyen de deux points :

voilà alors hein ^ on ne: ++ il est il est difficile d'aborder la: + la question du métier sans parler un petit peu des origines

¹ Développé par Robert Espesser, ainsi que l'éditeur MES.

3.2. Séquences de pauses remplies

Dans 12% des cas, les pauses remplies se retrouvent sous la forme d'une séquence complexe, faisant intervenir plusieurs allongements, plusieurs *eu*, ou une combinaison mixte d'allongements et de *eu* (éventuellement entrecoupés par une ou plusieurs pauses silencieuses). L'extrait ci-dessous montre un exemple de combinaison particulièrement longue (on notera que la suite se continue par une amorce, suivie à nouveau d'une pause remplie) :

eu Beaune est une **eu** la : la : la : **eu** le : cé- le :
cépage de : *eu* la ville de Beaune je veux dire

Au total, on observe 679 pauses remplies, qui constituent 591 séquences distinctes, dont le Tableau 2 donne la répartition quantitative.

Tableau 2. Types de pauses remplies

Type	N	sous-type	N
allongements	230	<i>simples</i>	216
		<i>complexes</i>	14
<i>eu</i>	323	<i>simples</i>	305
		<i>complexes</i>	18
mixtes	38		
Total	591		

4. ÉTUDE DES INTERACTIONS

Pauses silencieuses et pauses remplies peuvent difficilement être étudiées de façon indépendante :

- 380 des 1163 pauses silencieuses de notre corpus interviennent au contact d'une pause remplie (33%).
- Inversement, 344 des 591 séquences de pauses remplies sont au contact d'une pause silencieuse ou en contiennent une (58%).

C'est dire l'importance des interactions entre ces deux types de pauses.

4.1. Deux types de pauses remplies

Il nous semble important de distinguer du point de vue fonctionnel deux types de pauses remplies (ou de séquences). Les pauses remplies internes à un segment discursif constituent la majorité des cas (478 séquences sur 591, soit 81%). Elles marquent une interruption, suivie ou non d'une reprise et/ou d'une réparation :

eu Beaune est une **eu** la : la : la : **eu** le : cé- le :
cépage de : *eu* la ville de Beaune je veux dire

Dans 113 cas sur 591 séquences, toutefois (19%), les pauses remplies interviennent au début d'un segment :

enfin bon voilà eu le dé*but de mon voyage ça a été ça
++ || **eu** et après bon ben après les choses se sont mises
en place hein

Dans de tels cas, il ne s'agit pas d'une interruption liée à une difficulté de mise en place lexicale ou syntaxique de la fin du segment. Il s'agit sans doute d'« occuper le terrain » en attendant de trouver une suite au discours, et éviter ainsi que l'interlocuteur prenne la parole. Ce phénomène est très fréquent, puisqu'il affecte 113 segments sur 822, soit 14%. Ces 113 cas se répartissent en :

- 97 séquences commençant par *eu* ;
- 16 séquences commençant par un allongement.

Dans ce dernier cas, il s'agit (du moins dans notre corpus), uniquement de mots monosyllabiques et principalement des joncteurs, tels que *et*, *mais*, *puis* (11 cas sur 16). Les autres cas sont des mots-outils introducteurs de syntagmes (*de*, *où*, *le*).

4.2. Rôle des pauses silencieuses

Nous avons vu plus haut que les pauses silencieuses brèves ne sont jamais démarcatives. Parmi les 318 pauses silencieuses non démarcatives restantes, 289 (91%) sont associées à une pause remplie. Dans la grande majorité des cas (257 sur 289 soit 89%), une pause remplie précède la pause silencieuse :

on l'ap- au départ on faisait eu le **le**: + *eu* le: ma*cérer
le: le: **le poulet**: + le poulet bien sûr qui est issu aussi de
la Bourgogne puisque **eu** + le poulet de Bresse

Dans seulement 32 cas (11%), il y a interruption abrupte et la pause remplie intervient juste après :

et il a apprécié + **eu** ce Corton blanc

Il conviendrait d'étudier plus en détail ces cas, qui semblent intervenir après un mouvement mélodique majeur et/ou un accent.

Les 29 cas de pauses silencieuses non démarcatives qui ne sont pas associés à une pause remplie sont intéressants à analyser plus finement. Leur répartition est la suivante :

1. Dans 13 cas, il y a hésitation du locuteur, mais la pause silencieuse est en fait associée à une autre *marque du travail de formulation* :

- amorce (8 cas)

et puis là on vit au jour le j- + au jour le jour

- répétition (5 cas)

donc s- nous avons un rôle de + de soutien de marché

2. Dans 11 autres cas, la pause silencieuse est associée à une *marque discursive* :

- onomatopée (*pf*, etc.) (2 cas) :

alors que moi ça me dérange pas du tout au contraire:
pssff + je trouve que chacun a ses limites

- particule (*ben*, *hein*, etc.) (3 cas) :

elle apprendra comme moi sur le tas hein + de toute façon

- accent de focalisation (6 cas) :

*j'avais + te*llement soif de cette liberté-là*

Ce dernier phénomène est assez intéressant et mériterait d'être étudié sur des données plus volumineuses. Nous pouvons faire l'hypothèse qu'il permet à la fois de renforcer la perception de l'accent et au locuteur de refaire le plein de sa capacité pulmonaire avant, ou éventuellement après, l'émission de l'accent.

3. Il reste seulement 5 pauses silencieuses sur les 1163 du corpus qui ne soient associés à aucun autre indice :

elle va se sentir dans un + endroit de confiance donc

Dans ces 5 cas, il y a cohésion syntaxique du segment (il y a même liaison dans l'exemple ci-dessus). Il n'y a, de plus, aucune discontinuité intonative au moment de la pause silencieuse. Il conviendrait de recueillir plus d'exemples et d'analyser très finement le phénomène au niveau acoustique et syntaxique, mais on peut dire qu'il n'y a pas de perception d'une hésitation dans les cas que nous avons observés. Ces pauses se comportent donc exactement comme les pauses brèves. Leur durée est d'ailleurs relativement peu élevée (200 à 470 ms).

5. CONCLUSION

Nous avons montré dans cette étude, à partir d'un sous-ensemble du *Corpus de référence du français parlé* d'environ une heure et faisant intervenir 10 locuteurs (5 hommes et 5 femmes), que la pause silencieuse n'a en elle-même jamais le rôle de marque d'hésitation ou de travail de formulation. Elle n'intervient dans cette fonction qu'associée à d'autres marques, principalement les pauses remplies (allongement syllabique et item quasi-lexical *eu*). D'autres indices apparaissent également en association avec les pauses silencieuses ou remplies, tels que les amorces, les répétitions, ou des items quasi-lexicaux comme *ben*, *hein*, *pff*, etc. Une étude globale de la totalité de ces phénomènes liés au travail de formulation reste à conduire.

BIBLIOGRAPHIE

- [1] H. C. Barik. On defining juncture pauses: a note on Boomer's « Hesitation and grammatical encoding ». *Language and Speech*, 11, 156:159, 1968.
- [2] D.S. Boomer. Hesitation and grammatical encoding. *Language and Speech*, 8, 148:158, 1965.
- [3] E. Campione. *Étiquetage semi-automatique de la prosodie dans les corpus oraux : algorithmes et méthodologie*. Thèse de doctorat, Université de Provence, Aix-en-Provence, 2001.
- [4] M. Candéa. *Contribution à l'étude des pauses silencieuses et des phénomènes dits « d'hésitation » en français oral spontané*. Thèse de doctorat, Université Paris III, 2000.
- [5] H. Clark & E. Clark. *Psychology and Language*. New York : Harcourt, Brace, Jovanovich, 1977.
- [6] A. Cutler. The recognition of spoken words with variable representations, *Proceedings of the ESCA Workshop on the Sound Patterns of Spontaneous Speech* (pp. 83-92). Aix-en-Provence, France, 1998.
- [7] D. Duez. Salient pauses and non salient pauses in three speech style. *Language and Speech*, 25(7), 11:28, 1982.
- [8] D. Duez. The aim of SPoSS, *Proceedings of the ESCA Workshop on the Sound Patterns of Spontaneous Speech* (pp.VII-IX). Aix-en-Provence, France, 1998.
- [9] V. A. Fromkin. (1971). The non-anomalous nature of anomalous utterances. *Language*, 47, 27:52, 1971.
- [10] M. Goto, K. Itou & S. A. Hayamizu. A Real-time Filled Pause Detection System for Spontaneous Speech Recognition, *Proceedings of the 6th European Conference on Speech Communication and Technology (Eurospeech '99)* (pp. 227-230). Budapest, 1999.
- [11] I. Guaïtella. *Rythme et parole : comparaison critique du rythme de la lecture oralisée et de la parole spontanée*. Thèse de doctorat, Université de Provence, Aix-en-Provence, 1991.
- [12] A. E. Hieke, S. Kowal & D. C. O'Connell. The trouble with "articulatory" pauses. *Language and Speech*, 26, 203:214, 1983.
- [13] W. J. M. Levelt. *Speaking: From Intention to Articulation*. Cambridge MA : MIT Press, 1989.
- [14] M. A. Morel & L. Danon-Boileau. *Grammaire de l'intonation. L'exemple du français*. Paris : Ophrys, 1998.
- [15] F. C. M. Quimbo, T. Kawahara & S. Doshita. Prosodic analysis of fillers and self-repair in Japanese speech, *Proceedings of the International Conference on Spoken Language Processing (ICSLP)*. Sydney, Australia, 1998.
- [16] J. Vaissière. Rythm, accentuation and final lengthening in French. In J. Sundberg & L. Nord & R. Carlson (Eds.), *Music, Language, Speech and Brain*. Macmillan Press, 1991.
- [17] B. Zellner. *Caractérisation et prédiction du débit de parole en français. Une étude de cas*. Thèse de doctorat, Université de Lausanne, Lausanne, 1998.

