

Changements intonatifs dans la parole Lombard : au-delà de l'étendue de F_0

Pauline Welby

Institut de la Communication Parlée, CNRS UMR 5009, INPG, U. Stendhal, 46 av. F. Viallet, 38031 Grenoble, France
welby@icp.inpg.fr
http://www.icp.inpg.fr/~welby

ABSTRACT

Earlier studies of speech in noise (Lombard speech) have generally reported an increase in fundamental frequency (F_0). This study examined other potential intonational differences. Seven French speakers read a corpus of short paragraphs, in quiet and in 80 dB white noise. Four speakers increased F_0 range in noise. Six upscaled individual tones, although there was great inter-speaker variability. Noise had no effect on intonation pattern type; in particular, there was no tendency to produce more “early rises” in noise, even though these rises are cues to word segmentation. Producing an early rise (thus a LHLH or LHH pattern) may not add to the salience of the commonly produced LH pattern. There were no differences in tonal alignment, in contrast to earlier findings. This null result may be due to paradigm differences between the two experiments.

1. INTRODUCTION

1.1. L'effet Lombard

De nombreuses conversations ont lieu dans du bruit – celui des enfants qui jouent, des voitures qui passent, du vent qui souffle, pour n'en mentionner que quelques-uns. Pourtant, les locuteurs s'adaptent – ils parlent plus fort, modifient la durée des segments de parole (en augmentant la durée des voyelles par rapport à celle des consonnes, en général) et augmentent la fréquence fondamentale (F_0) (voir [8, 10, 14] sur ces modifications et d'autres). On appelle ces modifications l'effet ou le réflexe Lombard. Certaines au moins de ces modifications rendent la parole plus intelligible, aidant ainsi les auditeurs et permettant de transmettre plus efficacement le message [8, 14].

Quelques études ont examiné l'effet du bruit sur le F_0 , un effet noté en premier lieu par Lombard lui-même. En général, ces études notent une augmentation de F_0 et une grande variabilité inter-locuteurs (e.g., [8, 14]).

Très peu d'études, néanmoins, ont examiné l'influence du bruit sur les détails de la structure intonative. En fait, en général, les travaux qui ont traité des modifications de F_0 ont utilisé des listes de mots isolés (monosyllabiques pour la plupart), qui ne permettent pas d'examiner la structure intonative. Une étude fait exception, en ayant modélisé l'étendue tonale des cibles de F_0 (la hauteur relative des pics et des creux, le « tonal scaling ») de la parole néerlandaise [12].¹ Dans le bruit, tous les locuteurs ont

augmenté leur F_0 pour toutes les cibles. La fonction d'augmentation pourrait être décrite avec un seul modèle pour toutes les cibles, exceptée la dernière cible basse. Un des objectifs de l'étude présentée est d'examiner l'étendue tonale des cibles intonatives du français.

1.2. L'intonation du français

Tous les modèles de l'intonation du français s'accordent sur le fait que l'énoncé est divisé en unités plus petites, appelées Rhythmic Phrase [4], Accentual Phrase (syntagme accentuel, SA) [7], etc. Ce SA est caractérisé par une *montée finale* de F_0 sur sa dernière syllabe (s'il est non final dans l'énoncé) et une *montée initiale* facultative dans sa partie initiale. Des exemples sont donnés sur la figure 1, avec des transcriptions issues de la théorie autosegmentale-métrique, dans laquelle les patrons intonatifs sont considérés comme une série de tons bas ((L)ow) et hauts ((H)igh), les valeurs intermédiaires étant déterminées par interpolation [11]. Outre les patrons LHLH et LLH, un patron possible, souvent observé dans les SA courts, est LH, une montée de L1 à H2 (sans H1 et L2). Le début de la montée initiale (L1) est aligné avec le

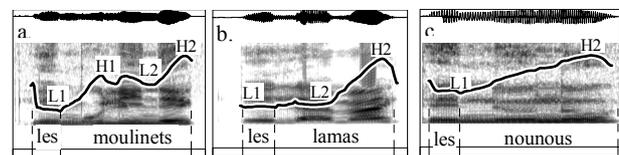


Figure 1 : Des SA avec (a) une montée initiale et une montée finale (LHLH), (b) une montée finale (LLH), (c) une montée d'un L initial à un H final (LH). Exemples tirés de l'étude.

début du premier mot lexical du SA (voir figures 1a, c) [15,17], formant souvent un « coude » dans la courbe de F_0 quand un mot outil précède le mot lexical. On observe ce coude souvent, même en l'absence de montée subséquente. Une montée initiale ou même un simple coude de F_0 (non suivi d'une montée) peuvent être utilisés pour détecter le début d'un mot lexical [16, 17].

1.3. Les changements intonatifs dans la parole Lombard en français

Une étude récente a examiné les modifications articulatoires et intonatives produites par une locutrice française dans le bruit [2, 18]. La parole a été enregistrée dans une condition silencieuse, en présence de bruit « cocktail party » à 85 dB et en présence de bruit blanc à

pitch change obtains because speakers deliberately raise their pitch.... We recognize, however, that some portion of the changes will also occur as a by-product of an increase in vocal effort (the “Lombard reflex”...), » mais les motivations pour cette revendication ne sont pas décrites.

¹ Plus précisément, les auteurs ne pensent pas que la comparaison porte sur la parole Lombard : « We assume that a considerable portion of the

85 dB. Trois différences intonatives ont été observées.

Montées initiales

Nous avons émis l'hypothèse selon laquelle les locuteurs produiraient plus de montées initiales dans le bruit, puisque ces montées devraient rendre la parole plus facile à segmenter et donc plus intelligible. Les résultats vont dans le sens de l'hypothèse – la locutrice a produit plus de SA avec des montées initiales dans le bruit. Cependant, il faut interpréter ces résultats avec précaution, d'abord parce qu'ils ne portent que sur une seule locutrice, et ensuite parce que les SA étaient de durée plus longue dans les conditions bruyantes que dans la condition de contrôle silencieuse. Cette différence de durée n'est pas inattendue, étant donné certains résultats de la littérature, mais elle complique l'analyse, puisqu'un SA plus long favorise la réalisation de la montée initiale [15, 17].

L'étendue tonale au sein du SA

La locutrice a augmenté son étendue de F_0 globale de la condition de contrôle silencieuse aux conditions bruyantes. Mais elle a également produit un patron d'étendue tonale inattendu : dans les SA avec un patron à 2 pics (LHLH), la chute du pic de la montée initiale (H1) au début de la montée finale (L2) était beaucoup plus faible dans le bruit que dans le silence. En fait, le L2 avait souvent une valeur de F_0 proche de celle du H1 qui le précédait, comme dans la figure 2b.



Figure 2 : Schéma des différences d'étendue tonale dans des patrons LHLH. (a) chute de H1 à L2, (b) plateau de H1 à L2.

Les patrons des figures 2a et b représentent les extrêmes d'un continuum, pas une dichotomie.² Pourtant, les patrons comme celui de la figure 2b sont plus fréquents dans les SA plus courts, pour lesquels les locuteurs n'arrivent pas à atteindre la cible L2 (voir [17], pp. 77–80). Le résultat de [18] est surprenant parce que la locutrice a produit ce patron (« LHplateauLH ») dans certains des SA plus longs.

L'alignement tonal

Les montées initiales produites dans le bruit ne commencent pas toujours à la frontière mot outil/ mot lexical, mais de façon surprenante, parfois au début du mot outil. Cet alignement inhabituel est peut-être dû à l'absence de *feed-back* auditif dans le bruit [5].

1.4. Objectifs de l'étude

L'objectif de cette étude est d'examiner l'influence du bruit sur trois aspects de la structure intonative soulevés par l'étude antérieure : le choix du patron intonatif, l'étendue des tons individuels et l'alignement de ces tons.

2. MÉTHODES

2.1. Corpus

Un corpus de 7 textes courts, chacun contenant 2 à 5 mots

cibles, a été construit. Il y avait 11 mots cibles à 2 syllabes et 11 à 3 syllabes. Les cibles ne contenaient que des syllabes CV avec des consonnes sonantes (soit des nasales, soit la latérale /l/). Cette contrainte segmentale a minimisé les perturbations segmentales de la courbe de F_0 . Chaque mot cible apparaissait en tête de phrase, précédé d'un article (*le, la, les* ou *un*). Ce syntagme était suivi d'une proposition relative commençant par un pronom relatif (*qui, que, qu'il*), suivie du syntagme verbal de la proposition principale. Un exemple avec cibles soulignées est présenté ci-dessous (1).

(1) Il y eut une scène chaotique à la crèche cet après-midi. Un mulot qui s'était réfugié au fond de la cantine avait fait peur aux petits. Ils s'étaient tous mis sur leurs chaises en hurlant. Les moulinets que Daniel était en train de faire avec les bras avaient fait tomber le bocal à poissons. Il y avait de l'eau et du verre partout. Heureusement, Charlotte pensa à mettre le pauvre petit poisson rouge dans un verre d'eau. Les nounous qui étaient venues chercher les enfants secouaient la tête en regardant la scène.

Chaque participante était assise à une table en face de la chercheuse et reçut la consigne « lisez les paragraphes comme si vous les lisiez à la personne en face de vous ».

2.2. Participantes

Sept locutrices natives du français hexagonal ont participé à l'étude. La Locutrice 7 était la locutrice de [18].

2.3. Procédures

Les participantes ont lu le corpus 4 fois, d'abord dans le silence avec un débit de parole normal puis avec un débit rapide, ensuite dans du bruit avec les deux débits. Du bruit blanc a été diffusé par écouteurs à 80 dB. Les locutrices ont été enregistrées avec un microphone-casque et un enregistreur numérique à 22,05 kHz, et les données ont été transférées sur ordinateur.

Les deux débits ont été utilisés uniquement pour obtenir de la parole avec des durées comparables pour les conditions de silence et de bruit, afin d'éviter la complication de l'analyse de l'étude précédente [18]. Nous espérons que la tendance à ralentir en lisant dans le bruit serait mitigée par le fait que la condition à débit normal bruyante suivait directement la condition à débit rapide silencieuse.

Les fichiers audio ont été segmentés et chaque énoncé sauvegardé comme un fichier individuel. Les courbes de F_0 et les spectrogrammes ont été créés sous Praat [3]. Les frontières de mots ont été étiquetées, en s'appuyant sur le signal et le spectrogramme, et à l'aide de scripts Praat qui ont permis de semi-automatiser le processus. (Pour une description de la procédure d'étiquetage, voir [15, 17], <http://www.icp.inpg.fr/~welby/praat.html>.)

3. RÉSULTATS

3.1. Variation de débit

De façon générale, la manipulation du débit n'a pas permis de minimiser les différences de durée entre la condition à débit normal silencieuse et la condition à débit normal bruyante. Une seule locutrice, la Locutrice 3, n'a pas produit de différence de durée significative entre ces

² Cette répartition va à l'encontre des modèles dans lesquels ces 2 contours (avec soit une chute de H1 à L2, soit un plateau) sont deux éléments distincts dans un inventaire de patrons holistiques (e.g., [13]).

deux conditions ($t(21) = 0,110; p = 0,914$). Une deuxième locutrice (la Locutrice 6) a des SA plus longs dans la condition à débit normal bruyante que dans la condition à débit normal silencieuse ($t(21) = -4,379; p < 0,001$). Les autres locutrices ont des durées plus courtes dans la condition à débit normal bruyante (toutes les valeurs de $p < 0,01$, à l'exception de la Locutrice 4, $p < 0,05$).

3.2. Patrons intonatifs observés

Contrairement aux prédictions, aucune locutrice n'a augmenté significativement le taux de montées initiales de la condition silencieuse à la condition bruyante.

Presque 90 % des mots cibles à 3 syllabes ont été produits avec une montée initiale, la plupart avec le patron LHLH, quelques-uns avec le patron LHH.

Pour les mots cibles à 2 syllabes, on a observé plus de variété dans les patrons intonatifs. Seuls 15 % des SA dans les conditions à débit normal silencieuse et à débit normal bruyante ont été produits avec soit LHLH soit LHH, tandis que 84 % ont été produits avec soit LH soit LLH. Trois locutrices sur 7 ont semblé préférer le patron LH, mais 3 locutrices ont produit de nombreux patrons LLH. La Locutrice 7 a produit de nombreux patrons LLH, mais aussi quelques patrons LHLH.

3.3. Alignement tonal

Nous avons inspecté visuellement l'alignement tonal du début de la montée initiale (L1) dans les conditions à débit normal silencieuse et bruyante. Les SA courts avec le patron LH n'ont souvent pas de coude, et la montée peut commencer tout au début du SA. Des SA plus longs avec

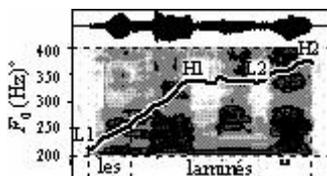


Figure 3 : Le SA *les laminés* réalisé avec le patron LHLH. (de la condition à débit normal bruyante de la Locutrice 6).

le patron LHLH, cependant, ont typiquement de L s coude L1. Contrairement à l'étude précédente, nous n'avons pas observé pour ces SA un alignement de L1 moins précis dans le bruit ; la plupart de ces SA a des coude L1 à la frontière mot outil/ mot lexical. Néanmoins, nous avons observé quelques SA LHLH dans lesquels L1 était réalisé au début du mot outil (voir la figure 3).

3.4. Étendue tonale

L'étendue de F_0 sur le SA cible ($H2 - 1$) a été examinée ; quatre locutrices ont augmenté leur étendue de F_0 du silence au bruit, comme le montre la figure 4.³ Des analyses de variance à 1-facteur ont été réalisées pour

³ Comme un des relecteurs l'a suggéré, il serait intéressant de voir si l'augmentation de F_0 est corrélée avec l'augmentation d'intensité typiquement observée dans le bruit. Malheureusement, une telle mesure n'est pas possible pour ce corpus, la distance microphone-bouche n'étant pas constante. (La mise en place des écouteurs pour la condition

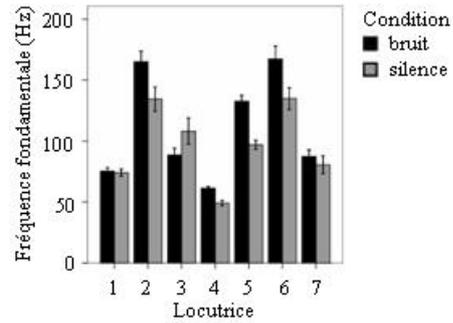


Figure 4 : Étendue de F_0 dans des SA cibles par condition de bruit. Les barres indiquent l'erreur standard de la moyenne.

déterminer si la condition de bruit a influencé l'étendue des 4 tons du patron LHLH, le plus fréquent des patrons. Les résultats sont donnés dans la table 1.⁴

Table 1 : Valeurs de F_0 pour les 4 tons du patron LHLH par condition. Les différences significatives sont marquées par '*'.

S	L1		H1		L2		H2	
	sil.	bruit	sil.	bruit	sil.	bruit	sil.	bruit
1	197	198	253	257	235	249	265	281
2	203	217*	257	272	251	290*	316	393*
3	193	199	270	260	214	216	279	274
5	208	227*	304	337*	260	292*	309	355*
6	216	231	305	331*	272	296*	334	373
7	179	186*	248	242	201	221*	255	265

Seules les Locutrices 2 et 7 (la locutrice dans [18]) ont produit des patrons proches du « LHplateauLH » observé dans [18]. Dans le bruit, elles ont augmenté L2 mais pas H1, ce qui a conduit à une plus petite chute de H1 à L2.

4. DISCUSSION ET CONCLUSIONS

Parmi les paramètres examinés, les résultats se sont avérés les plus fructueux pour l'étendue tonale. Pour toutes les locutrices sauf une, nous avons constaté des différences significatives dans l'étendue des tons individuels entre les conditions de silence et de bruit, même s'il y avait une variabilité inter-locutrices importante en ce qui concerne quel(s) ton(s) étai(en)t impliqué(s). Cette variabilité n'est pas surprenante : certains travaux sur le français ont montré de grandes différences inter-locuteurs pour l'étendue [6, 19]. Le fait que l'on trouve des différences d'étendue des tons individuels est d'autant plus remarquable que toutes les locutrices n'ont pas augmenté leur étendue de F_0 globale dans la condition bruyante (un résultat qui s'accorde avec la variabilité observée dans la littérature). Une mesure de l'étendue globale de F_0 n'est donc pas adaptée à ce type de différence.

Les différence entre les résultats de notre étude, dans laquelle 7 locutrices françaises ont montré une grande variabilité dans l'étendue des tons individuels, et ceux de

de bruit a souvent nécessité un repositionnement du microphone.) Nous remarquons, néanmoins, qu'une augmentation de F_0 n'est pas systématiquement observée, ni pour les locutrices de cette étude, ni dans d'autres études (contrairement à une augmentation d'intensité).

⁴ Les données de la Locutrice 4 ont été exclues, puisqu'elle a produit trop peu (3) de patrons LHLH dans la condition bruyante.

[12], dans laquelle 15 locuteurs néerlandais ont systématiquement déplacé tous les tons vers le haut, sont peut-être linguo-spécifiques. L'étendue tonale varie d'une langue à une autre (e.g., le français vs. le néerlandais et le grec [1]) pour la parole produite dans des conditions claires. C'est peut-être aussi le cas de la parole Lombard. Puisque la forme exacte du F_0 du SA ne signale pas, en général, une différence de sens en français (au moins pour les SA non ambigus), il semble possible que les locuteurs n'aient pas besoin de préserver toutes les caractéristiques de l'étendue tonale.

Il est aussi possible que ces différences soient dues à des différences de paradigmes expérimentaux, puisque des facteurs telles que l'intensité et le type de bruit pourraient influencer l'étendue tonale.

L'étude présentée ici n'a pas montré un alignement moins précis de L1 dans le bruit, même pour la Locutrice 7 (la locutrice dans [18]). Ce résultat contraire à celui de [18] est peut-être dû à des différences de paradigmes expérimentaux. L'étude antérieure a utilisé du bruit à 85 dB diffusé par une enceinte (ce qui a mis la locutrice comme les chercheuses/ auditrices dans un environnement bruyant), alors que dans l'étude présentée ici, le bruit était à 80 dB et diffusé par des écouteurs. Il est possible que l'alignement tonal ne soit perturbé qu'au-delà d'un certain seuil ou avec un certain degré d'effort vocal.

Finalement, les locutrices n'ont pas produit plus de montées initiales dans le bruit. Ce résultat est peut-être dû au fait que pour la plupart des locutrices, les SA ont été produits avec une durée moins longue dans le bruit, ce qui défavorise la réalisation de la montée initiale. Cependant, nous n'avons pas constaté d'augmentation de taux de montées initiales, même pour les deux locutrices dont les SA n'étaient pas de durée plus courte dans les conditions bruyantes. Il se pourrait que la montée du patron LH, que l'on observe souvent dans des SA courts, soit aussi efficace comme indice de la segmentation lexicale que la montée initiale du LHLH. Plus précisément, alors que l'ajout d'une montée initiale à un patron LLH (donnant LHLH) pourrait rendre plus saillant le début d'un mot lexical, il se pourrait qu'il n'existe aucun avantage pour LHLH versus LH.

De nombreux auteurs ont noté l'influence de la tâche dans l'étude de la parole Lombard (e.g., [9]). En particulier, augmenter la charge communicative (en utilisant des paires locuteur-auditeur, par exemple) pourrait conduire à une accentuation de certains aspects de l'effet.

La plupart des études Lombard ont supposé au moins tacitement que certains des changements sont indépendants de la langue considérée, comme le terme « *réflexe Lombard* » le suggère. Même si l'on ne peut pas conclure des résultats de notre étude qu'il existe des indices linguo-spécifiques, cette question mérite d'être examinée.

BIBLIOGRAPHIE

[1] A. Arvaniti, D. R. Ladd et I. Mennen. Stability of tonal alignment: the case of Greek prenuclear accents. *J. Phon.*, 26 : 3–25, 1998.

- [2] L. Bailly, *Étude articulatoire de la parole produite en environnement bruyant*. Mémoire de master. Laboratoire d'Acoustique Musicale, Université de Paris 6, 2005. http://www.icp.inpg.fr/~lbailly/RapportFinal_LucieBailly.pdf
- [3] P. Boersma et D. Weenink. PRAAT: Doing phonetics by computer (version 4.3.28), 2005. <http://www.praat.org>
- [4] A. Di Cristo. Intonation in French. Dans D. Hirst et A. Di Cristo (éds.), *Intonation Systems: a survey of twenty languages*, pages 195–218. Cambridge : Cambridge University Press, 1998.
- [5] L. Elliott et A. Niemoeller. The role of hearing in controlling voice fundamental frequency. *Intl. Audiology*, 9 : 47–52, 1970.
- [6] C. Fougeron et S.-A. Jun. Rate effects on French intonation: Prosodic organization and phonetic realization. *J. Phon.*, 26 : 45–69, 1998.
- [7] S.-A. Jun et C. Fougeron. Realizations of accentual phrase in French. *Probus*, 14 : 147–172, 2002.
- [8] J.-C. Junqua. The Lombard reflex and its role on human listeners and automatic speech recognizers. *JASA*, 93 : 510–524, 1993.
- [9] H. Lane et T. Bernard. The Lombard sign and the role of hearing in speech. *J. Speech & Hearing Research*, 14 : 677–709, 1971.
- [10] É. Lombard. Le signe de l'élévation de la voix. *Annales des maladies de l'oreille et du larynx*, 37 : 101–119, 1911.
- [11] J. B. Pierrehumbert. *The phonology and phonetics of English intonation*. Thèse de doctorat, MIT, 1980.
- [12] E. Shriberg, D. R. Ladd, J. Terken et A. Stolcke. Modeling pitch range variation within and across speakers: Predicting F0 targets when “speaking up.” *Proc. ICSLP 4*, pages 1–4 (annexe des actes), 1996. Version corrigée disponible à <http://www.asel.udel.edu/icslp/cdrom/vol2/553/a553.pdf>
- [13] J. Vaissière. Une procédure de segmentation automatique de la parole en mots prosodiques en français. *Actes des VIIèmes JEP*, pages 103–114, 1977.
- [14] W. Van Summers, D. B. Pisoni, R. H. Bernacki, R. I. Pedlow et M. A. Stokes. Effects of noise on speech production: acoustic and perceptual analyses. *JASA*, 84 : 917–928, 1988.
- [15] P. Welby, French intonational structure: evidence from tonal alignment. *J. Phon.*, sous presse.
- [16] P. Welby. French intonational rises and their role in speech segmentation. *Proc. Eurospeech 8*, pages 2125–2128, 2003.
- [17] P. Welby. *The slaying of Lady Mondegreen, being a study of French tonal alignment and association and their role in speech segmentation*. Thèse de doctorat, The Ohio State University, 2003. <http://www.ling.ohio-state.edu/publications/dissertations>
- [18] P. Welby, L. Bailly, M. Garnier, M. Dohen et H. Løevenbruck. Pitch changes in noisy conditions: data from French. Dans *Proc. Making Europe More Attractive for Researchers*, à par.
- [19] P. Welby et H. Løevenbruck. Anchored down in Anchorage: syllable structure and segmental anchoring in French. *Italian Journal of Linguistics. Special issue on Autosegmental-metrical approaches to intonation in Europe: Tonal targets and anchors*, M. D'Imperio (éd), à par.

REMERCIEMENTS

Je tiens à remercier Hélène Løevenbruck pour ses conseils précieux et pour son aide avec la construction du corpus, et Denis Beauteemps, Xavier Pelorson et Coriandre Vilain pour leur assistance technique. Je remercie également les participantes. Cette recherche a été soutenue par une bourse internationale Marie Curie du 6th European Community Framework Programme.