



L'emphase en anglais : critères acoustiques et analyse statistique. Tentative de définition

Sophie Herment-Dujardin, Daniel Hirst

► To cite this version:

Sophie Herment-Dujardin, Daniel Hirst. L'emphase en anglais : critères acoustiques et analyse statistique. Tentative de définition. Interfaces Prosodiques, Mar 2003, Nantes, France. pp.51-56, 2003. <hal-00369933>

HAL Id: hal-00369933

<https://hal.archives-ouvertes.fr/hal-00369933>

Submitted on 23 Mar 2009

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.

L'emphase en anglais : critères acoustiques et analyse statistique. Tentative de définition.

Sophie Herment-Dujardin* & Daniel Hirst**

Abstract

This mainly experimental study aims at determining the most relevant acoustic and prosodic parameters of emphasis in English spontaneous speech. A perception test based on naïve native speakers was carried out, making it possible to determine a perceived degree of emphasis for each word. The corpus was manually aligned for both phonemes and rhythmic units. An acoustic analysis was then carried out, based on the hypothesis that fundamental frequency, segmental duration, global intensity and pauses may each contribute to the perception of emphasis. A regression-tree analysis was applied both to the complete corpus and also to each speaker individually as well as to each type of discourse, in order to bring to light both general tendencies and a number of speaker-specific strategies.

1. Introduction

Cette étude, fondée sur le discours spontané anglais, a pour but de tenter de définir le phénomène d'emphase. En effet de nombreux travaux font référence à l'emphase, cette notion semble acquise pour tous et pourtant, le terme n'est jamais défini au préalable. Nous sommes donc partis d'une notion volontairement vague de ce concept afin de tenter de le définir. Notre réflexion se situe dans un cadre phonologique et phonétique et c'est en analysant les critères acoustiques et prosodiques du phénomène que nous avons tenté de trouver des réponses à nos questions : qu'est-ce qui fait que nous percevons dans le discours des mises en relief, des éléments plus proéminents que d'autres, bref, des emphases ?

Nous décrivons ici notre étude, qui s'appuie sur des travaux antérieurs en ce qui concerne les paramètres choisis pour l'analyse. Nous avons opté pour une étude perceptive afin de déterminer quels étaient les passages emphatiques de notre corpus. Nous avons procédé à l'analyse acoustique de ces segments puis à une analyse statistique.

2. Base de données et méthode

Notre travail est fondé sur une base de données de parole spontanée en anglais. Elle regroupe trois types de discours : du débat politique, de la conversation informelle et une émission de radio où deux femmes traitent d'un sujet émotionnel qui les touche particulièrement. Nous avons sélectionné certains passages de ces enregistrements, qui constituent notre corpus primaire, sans prendre les morceaux où plusieurs personnes parlent en même temps par exemple.

Afin de déterminer quels étaient les passages ou mots emphatiques de notre corpus primaire, nous avons opté pour une étude perceptive : nous avons réalisé des sondages auprès d'auditeurs anglophones naïfs, leur demandant de pointer les passages ou mots qui leur semblaient emphatiques dans les morceaux écoutés. Grâce à ces sondages, nous avons pu déterminer un degré d'emphase pour chaque mot, fondé sur le pourcentage d'auditeurs ayant pointé le mot. Cet « indice d'emphase » a servi de point de départ à nos analyses acoustiques. Cette méthode nous semblait la mieux adaptée, étant la plus objective. Nous avons en effet voulu nous baser sur ce qui est intuitivement perçu par des auditeurs naïfs et non imposer une idée préconçue de ce qu'est l'emphase en évitant ainsi une démarche circulaire. Comme indiqué ci-dessus, nous sommes partis d'une définition volontairement vague de l'emphase.

* Laboratoire Parole et Langage, CNRS UMR 6057, Université de Provence. dujardin@up.univ-aix.fr

** Laboratoire Parole et Langage, CNRS UMR 6057, Université de Provence. daniel.hirst@up.univ-aix.fr

3. Analyse acoustique

Nous avons entrepris de rechercher les paramètres les plus pertinents dans l'emphase. La fréquence fondamentale et la durée sont généralement reconnues comme les corrélats essentiels de l'emphase. Cooper *et al.* (1985) et Eady *et al.* (1986) ont montré qu'à la fois la fréquence fondamentale et la durée augmentent sur les occurrences emphatiques mais que la relativité est à prendre en compte aussi : les mots adjacents sont importants et c'est pourquoi il semble essentiel de mesurer chaque mot par rapport aux mots qui l'entourent. Tannen (1984) a étudié "high-involvement style", le style marquant une forte implication, et donne comme critères pertinents la hauteur mélodique, l'amplitude, la qualité de la voix et les pauses. Hirst & Di Cristo (1998:32), qui passent en revue les systèmes intonatifs de vingt langues, affirment que l'emphase est manifestée principalement par des mouvements mélodiques plus amples souvent accompagnés d'une intensité plus forte et d'une durée plus importante¹.

Nous avons donc effectué une analyse acoustique de nos segments, en émettant l'hypothèse que quatre critères étaient certainement essentiels dans la perception de l'emphase : la fréquence fondamentale, la durée segmentale, l'intensité globale et les pauses. Nous avons choisi ces quatre critères au vu des études antérieures.

Cette analyse acoustique, réalisée grâce au logiciel PRAAT (Boersma, P. & Weenik, D., 1996), nous a permis dans un premier temps de vérifier nos hypothèses et de dégager ensuite des paramètres qui nous paraissaient importants, dans le but de mettre en œuvre une analyse statistique.

3.1. Fréquence fondamentale

Trois niveaux rythmiques ont été dégagés pour les analyses : le mot (words), l'unité tonale², (UT) et l'unité intonative (UI) (cf. figure 1) :

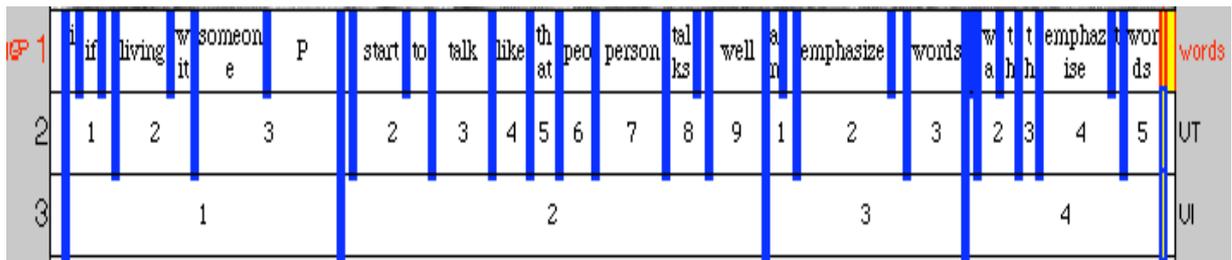


Figure 1 : P2.6S08 : TextGrid avec trois "tiers" : words, UT, UI.

Pour chaque niveau ont été extraites les valeurs maximales et minimales, ainsi que les écarts-type et les pentes. Les valeurs maximales et minimales permettent de voir si l'on se situe dans le grave ou dans l'aigu, la pente permet de voir si le contour intonatif est plat ou non et l'écart-type nous renseigne sur l'ampleur du mouvement mélodique.

3.2. Durée

Nous avons utilisé pour cette étude les données de Campbell (Campbell, N., 1992a & 1992b), qui permettent de prédire les durées moyennes des mots. La comparaison des durées prédites et des durées réelles nous permettra de voir s'il y a un allongement ou un raccourcissement des mots de notre corpus en fonction de l'emphase.

3.3. Intensité

L'intensité n'a pas été mesurée sur les mots, car elle ne peut être déterminée pour des segments inférieurs à 64 ms dans le logiciel PRAAT. L'intensité a donc été mesurée sur les unités tonales et intonatives. Nous avons extrait les valeurs minimales et maximales ainsi que les moyennes et les écarts-type à nouveau.

¹ "in the majority of languages described in this volume, focalisation and/or emphasis is said to be best manifested by an extra pitch prominence, giving rise to larger F0 movements often accompanied by extra intensity and duration" (Hirst & Di Cristo, 1998:32).

² Ce que nous appelons unité tonale, à la suite de Hirst (1999) et Hirst *et al.* (2000), est une unité quelque peu différente pourtant, totalement dépendante des frontières de mot pour nous, qui est constituée d'un mot accentué (et en inclut donc le cas échéant l'anacrusse) suivi de tous les mots inaccentués jusqu'au mot accentué suivant.

3.4. Pauses

Il semblait très ardu de trouver des paramètres pertinents concernant les pauses pour une analyse statistique. Les différentes analyses effectuées sur les pauses ne se sont pas révélées probantes (cf. Herment-Dujardin, 2001:159-176). Nous n'avons pas réussi à dégager un paramètre commun à tous les locuteurs et nous avons donc décidé de n'intégrer dans notre analyse statistique que le paramètre pause avant ou après un mot, sans les durées.

Les analyses effectuées sur les pauses ont cependant confirmé l'hypothèse que les locuteurs peuvent avoir différentes stratégies (cf. Dahan & Bernard, 1996, ou Gussenhoven & Rietveld, 1998, par exemple) pour exprimer l'emphase, en ce qui concerne les pauses, mais pas uniquement probablement (cf. 3.4).

3.5. Amplitude totale

Ce paramètre est inspiré de Beckman (1986) : nous avons multiplié pour chaque mot sa durée réelle et son intensité moyenne, ce qui correspond à une énergie (E), l'énergie dépensée pour produire un mot (cf. figure 2).

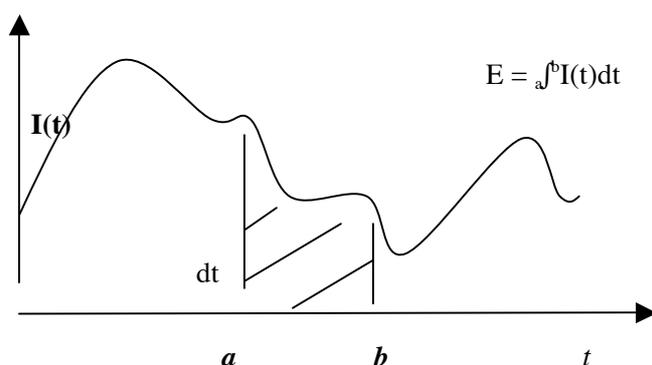


Figure 2 : amplitude totale

4. Analyse statistique

Grâce aux analyses acoustiques décrites ci-dessus nous avons dégagé des éléments caractéristiques et extrait des paramètres pertinents dont nous avons entrepris l'analyse statistique. Le logiciel CRUISE (Kim & Loh, 2001) a été utilisé pour cette étude.

4.1. L'emphase : phénomène catégoriel ou continu ?

Cruise est un logiciel très performant qui permet de faire des choix et de prendre des options assez différentes qu'il serait trop long de relater ici (cf Herment-Dujardin, 2001:184-234 pour les détails). Nous avons effectué de nombreux essais avec différentes options et ceux-ci nous ramènent à un point largement discuté par Ladd & Morton (1997) en particulier : l'emphase est-elle un phénomène catégoriel ou continu ? Nous avons opté pour un mélange de ces deux aspects : oui, l'emphase est un phénomène catégoriel, on peut dire qu'un élément est emphatique ou non. Mais à l'intérieur de l'emphatique, on peut voir du continu. Nous avons donc déterminé trois classes d'emphase :

- classe 0 : non emphatique
- classes 1 et 2 emphatiques, la classe 1 étant moins emphatique que la classe 2.

Nous avons effectué l'analyse statistique sur tout le corpus, puis par type de discours et enfin par locuteurs.

4.2. Corpus primaire

L'amplitude totale s'est révélée le paramètre le plus pertinent. Si cette dernière est moyenne, alors l'intensité et la F0 entrent en jeu : si un mot a une intensité plus forte ou est plus aigu que les mots qui l'entourent dans l'unité intonative, alors il va être perçu emphatique. Si ce n'est pas le cas mais qu'en contrepartie le mot est prononcé avec un mouvement mélodique important, il va également être perçu emphatique.

4.3. Types de discours

Les résultats ici ne sont pas très probants. Il semble que plus le nombre de locuteurs est élevé, plus on a de difficultés à parvenir à des prédictions correctes.

4.4. Locuteurs

Les résultats par locuteurs montrent que l'emphase est bien rendue différemment selon les locuteurs. Chez certains, l'énergie est déterminante, chez d'autres c'est la fréquence fondamentale : le mouvement mélodique souvent, ou bien la hauteur tout simplement.

5. Conclusion

L'analyse statistique entreprise grâce à un logiciel tout récent et extrêmement performant devait nous permettre d'extraire les paramètres les plus révélateurs de l'emphase sur 2375 occurrences et 11 locuteurs dans 3 types de discours différents.

Nous avons opté pour une notion dichotomique de l'emphase avec deux degrés dans l'emphatique.

La diversité du corpus et le nombre de locuteurs ont constitué une difficulté majeure en ce qui concerne les prédictions. Cela signifie qu'il existe plusieurs moyens d'exprimer l'emphase en anglais et ces moyens varient considérablement d'un locuteur à l'autre.

Grâce à la performance du logiciel CRUISE nous pouvons néanmoins dégager des constantes et affirmer que l'énergie est un paramètre fondamental dans la perception de l'emphase. Cela signifie que si un mot est allongé et prononcé avec une intensité plus forte que sur les mots adjacents, il sera généralement perçu emphatique. Si au contraire l'énergie dépensée est faible, le mot sera perçu non emphatique.

Lorsque l'énergie n'est pas déterminante, alors la hauteur mélodique le devient. Le mouvement mélodique sur le mot est important : s'il est ample et qu'il part de plus haut que pour les mots adjacents, le mot sera perçu emphatique.

Enfin l'intensité est déterminante aussi. Plus l'intensité est forte par rapport aux mots adjacents, plus le mot sera perçu emphatique.

N'oublions pas de nuancer ces affirmations. Ces règles sont générales et ne marchent pas à 100%. En effet, chaque locuteur semble avoir ses habitudes et ses façons propres d'exprimer l'emphase, ses stratégies personnelles.

Cette étude statistique confirme que l'emphase est un phénomène complexe mais que l'on peut tout de même énoncer certains principes généraux.

L'importance du sémantisme, de la syntaxe et du contexte ne sont certes pas à négliger (cf. Herment-Dujardin & Hirst, 2002), mais tel n'était pas notre propos ici. Nous nous en tenons à une définition en termes de critères acoustiques et prosodiques uniquement et nous proposons donc la définition suivante de l'emphase :

l'emphase est une mise en relief inattendue qui marque une implication de la part du locuteur plus importante qu'à la normale. Cette mise en relief est obtenue en général grâce à une augmentation de l'énergie dépensée pour prononcer l'occurrence emphatique, et/ou à une augmentation de la hauteur et/ou du mouvement mélodique sur cette occurrence.

6. Références

- Beckman, M., 1986, *Stress and Non-Stress Accent*, Dordrecht : Foris Publications.
- Boersma, P. & Weenik, D., 1996, PRAAT : a system for doing phonetics by computer, *Report of the Institute of Phonetic Sciences of the University of Amsterdam*, 132.
- Campbell, N., 1992a, Syllable-based segmental duration, in Bailly G. & Benoît C. (eds.), *Talking Machines*, Amsterdam : North Holland, p. 211-224.
- Campbell, N., 1992b, *Multi-level Timing in Speech*, PhD Thesis, Sussex University.
- Cooper, W.E., Eady, S.J., Mueller, P.R., 1985, Acoustical aspects of contrastive stress in question-answer contexts, *Journal of the Acoustical Society of America* 77/6, p. 2142-2156.

- Dahan, D. & Bernard, J-M, 1996, Interspeaker variability in emphatic accent production in French, *Language and Speech* 39/4, p. 341-374.
- Eady, S.J., Cooper, W.E., Klouda, G.V., Mueller, P.R., Lotts, D.W., 1986, Acoustical characteristics of sentential focus : narrow vs. broad and single vs. dual focus environments, *Language and Speech* 29/3, p. 233-251.
- Gussenhoven, C. & Rietveld, T., 1998, On the speaker-dependence of the perceived prominence of F0 peaks, *Journal of Phonetics*, 26, p. 371-380.
- Herment-Dujardin, S. & Hirst, D., 2002, Emphasis in English : a perceptual study based on manipulated synthesized speech, in *Proceedings of the 1st International Conference on Speech Prosody*, (Aix-en-Provence, 11-13 avril 2002), p. 379-382.
- Herment-Dujardin, S., 2001, L'emphase dans le discours spontané anglais : corrélats acoustiques et prosodiques, Thèse de Doctorat, Laboratoire Parole et Langage, Aix-en-Provence : Université de Provence.
- Hirst, D.J. & Di Cristo, A., 1998, *Intonation Systems : a Survey of Twenty Languages*, Cambridge : Cambridge University Press.
- Hirst, D.J., Di Cristo, A., Espesser, R., 2000, Levels of representation and levels of analysis for intonation, in M. Horne (ed), *Prosody, theory and experiment*, Dordrecht : Kluwer Academic Publishers, p. 51-88.
- Hirst, D.J., 1999, The symbolic coding of duration and timing : an extension to the INTSINT system, *Proceedings Eurospeech '99* (septembre 1999, Budapest), p.1639-1642.
- Kim, H. & Loh, W-Y, 2001, CRUISE, User Manual, *Technical Report 989*, March 3, 1998, revised November 10, 2001, Department of Statistics, University of Wisconsin, Madison.
- Ladd, R & Morton, R, 1997, The perception of intonational emphasis, continuous or categorical ?, *Journal of Phonetics*, 25, p. 313-342.
- Tannen D., 1984, *Conversational Style*, Norwood, NJ : Ablex.