

Intonation, wh-word questions et attitudes en anglais américain, Une étude acoustico-perceptive

Mohamed Abdelbassat MGHIZOU

Laboratoire de Phonétique et Phonologie (UMR 7018), CNRS / Sorbonne Nouvelle, Paris, France
mghizou_28@yahoo.co.uk

ABSTRACT

This paper aims at investigating fundamental frequency, duration and intensity of some American English wh-word interrogatives, expressed in neutral situation and attitudinal one. Three attitudes are taken into account: neutral, doubt and surprise. Declination Line theory as proposed by Maeda model [3], [4] is not always consistent with our data mainly when attitudes are expressed. F0 plays an important role in the distinction between the two attitudes under study whereas, duration plays a secondary role in this distinction. The results were confirmed by three perception tests, with manipulation done on natural speech and synthetic speech.

1. INTRODUCTION

La parole, le propre de l'homme, sert non seulement à transmettre des messages linguistiques via un nombre illimité de phrases, mais elle sert également à véhiculer l'état émotif et attitudinal que certains locuteurs veulent exprimer dans leur discours. Cette dernière fonction « complexe » de la parole a mené plusieurs chercheurs vaissière [10] à ne pas inclure la parole expressive dans leurs descriptions et particulièrement au niveau de la prosodie.

Cette étude traite un aspect de la prosodie, qui est l'intonation. Elle analyse la parole attitudinale exprimée dans l'interrogation en anglais américain. Elle vise à :

- Dégager les traits acoustiques qui différencient les trois attitudes en question en se basant sur trois paramètres principaux : f0, durée et intensité
- Déterminer le jeu de la ligne de déclinaison et des paramètres globaux éventuels dans la différenciation entre les trois types d'attitudes.
- Faire une corrélation entre l'intention des locuteurs et le jugement des auditeurs.

2. ANALYSE ACOUSTIQUE

2.1. Corpus :

Le corpus choisi dans cette étude se compose de huit phrases lues représentant les wh-word questions en anglais américain. Ces phrases ont été prononcées dans trois contextes différents : demande neutre, demande de doute et demande de surprise. Le tableau ci dessous présente une description détaillée du corpus sous l'étude :

Tableau 1 : nombre mots, syllabes et syllabes accentuées constituant le corpus sous l'étude.

Phrases	N de mots	N de syllabes	N de syllabes accentuées
1	4	6	2
2	8	12	3
3	7	8	2
4	7	11	3
5	6	9	3
6	9	12	3
7	10	11	2
8	8	10	2

Pour réaliser une consistance entre les différentes réalisations prosodiques, chaque phrase a été prononcée trois fois, résultant un total de 72 tokens. Il reste à signaler que deux américaines natives ont participé à cette expérience.

2.2. Procédure

L'enregistrement a été effectué dans la chambre sourde de ILPGA (Institut de Linguistique et de Phonétique Générale et appliquée, Sorbonne Nouvelle PARIS III-). Le matériel utilisé pour enregistrer est un MINIDISC, de type SONY, un appareil enregistreur de MINIDISC de type SONY également ainsi que deux microphones stéréo. Durant l'enregistrement, plusieurs précautions ont été prises afin d'avoir un signal de très bonne qualité. Les microphones ont été placés 20 centimètres environ loin de la bouche des sujets, en plus et pour avoir une meilleure représentation du signal, ces derniers ont été demandés de lire les phrases le plus fort que possible, et un essai de 5 minutes a été établi pour que les sujets puissent se familiariser avec le corpus. En fin, le signal a été numérisé avec le logiciel Sound Forge (version 6) à un taux d'échantillonnage de 22.000 HZ via 16 bits analogue to digital converter.

Toutes les répétitions ont été segmentées sur Praat, et l'analyse de f0, durée et intensité a été faite en utilisant les scripts programmés par Cédric Gendrot.

2.3. Résultats

L'analyse acoustique du f0, durée et intensité ont montré les résultats suivants :

Demande neutre

- F0 : un contour mélodique commençant généralement par une montée initiale très importante, suivie d'une descente sur les syllabes intermédiaires puis d'une

remontée sur la dernière ou les deux dernières syllabes pour marquer l'interrogation.

- **Durée** : des noyaux de dernières syllabes qui sont généralement les plus longs. Par contre, en ce qui concerne les syllabes, la durée de ces dernières varie selon le nombre des syllabes constituant la phrase. Mais généralement ce sont les premières et les dernières syllabes qui sont les plus allongées.

- **Intensité** : au fur et à mesure que les deux sujets procèdent dans leur réalisation des phrases, l'intensité de leurs syllabes baisse.

Demande de doute

- **F0** : un contour mélodique qui prend la première syllabe comme point de départ pour continuer de s'accroître jusqu'à la fin de la phrase. Ceci est contraire à ce qui était constaté pour la demande neutre. Ainsi, les valeurs moyennes de f0 dans la demande de doute sont plus importantes que pour celles de la demande neutre.

- **Durée** : première et une dernière syllabe de chaque énoncé qui sont les plus allongées, soit au niveau du noyau ou de la syllabe entière. Nous avons constaté également que la demande de doute est produite avec un débit plus rapide que la demande neutre.

- **Intensité** : nous avons constaté que les deux sujets ont une tendance générale à diminuer leur intensité en procédant dans la réalisation de la phrase. C'est-à-dire que les dernières syllabes sont généralement moins intenses que les premières.

Demande de surprise :

- **F0** : un même schéma mélodique pour la demande de surprise que pour la demande de doute. Sauf que le fondamental de cette dernière est généralement un petit plus bas par rapport à la demande de surprise.

- **Durée** : un allongement des noyaux de la première et la dernière syllabe de la part des deux sujets en exprimant la demande de surprise. En terme de syllabe, c'est celle du wh-word qui est généralement la plus allongée. De plus, la demande de surprise est réalisée par les deux sujets avec un débit plus lent par rapport à la demande de doute.

- **Intensité** : les dernières syllabes sont généralement moins intenses par rapport aux premières syllabes.

Le tableau ci dessous, illustre en chiffres les résultats présentés ci dessus.

Tableau 2 : valeurs moyennes de f0, durée et intensité des phrases interrogatives produites dans les trois situations (neutre, doute et surprise).

valeurs moyennes	F0		Durée	Intensité	
	syllabe Initial	syllabe Finale	Phrases ayant le même nombre de syllabes	syllabe Initial	syllabe Finale
Attitudes					
Demande neutre	193HZ	279HZ	1601ms	61dB	46dB
Demande de doute	251HZ	434HZ	1531ms	58HZ	47HZ
Demande de surprise	287HZ	465HZ	1733 ms	64dB	51dB

Comme nous l'avons signalé auparavant, la description des différentes réalisations prosodiques au niveau de la phrase nécessite l'établissement d'une construction linéaire tel que la Ligne de Base Maeda [3], [4]; Vaissière [9]. Il existe d'autres constructions capables de produire une telle description comme celle de Garding [1]; Thorsen, [7], mais nous avons opté pour celle de la Ligne de base. L'idée derrière ce choix est d'observer la trajectoire de la Ligne de Base durant les différentes réalisations et son rôle important dans la distinction entre les attitudes sous l'étude. Nous avons obtenu les figures suivantes :

Neutral

Doubt

Surprise

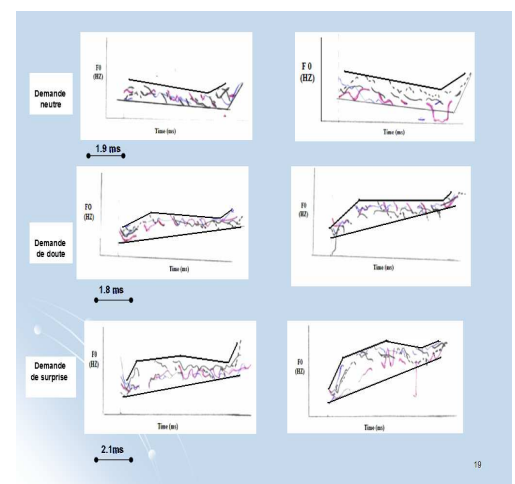


Figure 1: superposition des phrases wh-word interrogatives représentant les trois attitudes et ayant presque la même durée : les schémas à gauche appartiennent à S1 et les autres à droite à S2 (After Mghizou [5]).

Donc, nous pouvons parler d'une déclinaison de la ligne de base lors de la prononciation des phrases interrogatives représentant la demande neutre. Par ailleurs, la demande de doute et de surprise prend un format ascendant de la

ligne de base avec une distinction au niveau du fondamental qui est plus élevé dans le cas de la demande de surprise. Ainsi, cette dernière est produite avec un débit moins rapide que la demande de doute.

Pour corroborer nos résultats acoustiques concernant le rôle pivot du fondamental dans la distinction entre les trois attitudes, nous avons appliqué un test ANOVA. La variable indépendante était le type d'attitude, i.e. Neutre, doute et surprise; et la variable dépendante était les valeurs moyennes de f_0 mesurées sur le noyau de la dernière syllabe de la montée finale d'interrogation. Nous avons obtenu :

$F(2, 42) = 31,692$ et $P < 0,0001$, ce qui veut dire que la structure du fondamental joue un rôle très important dans la distinction entre les trois attitudes et :

f_0 surprise > f_0 doubt > f_0 neutral.

3. ANALYSE PERCEPTIVE

3.1. Procédure

Pour tester et vérifier la fiabilité de nos résultats acoustiques, nous avons eu recours à trois tests de perceptions précédés par une séance d'entraînement. Le premier, comprends 24 stimuli de la parole naturelle lue, il est pour but de vérifier si les auditeurs américains natifs (huit auditeurs) peuvent identifier et distinguer entre la demande neutre, la demande de doute et la demande de surprise. En ce qui concerne le deuxième et le troisième test, ils consistent de la parole synthétisée. Les mêmes stimuli présentés dans le premier test ont été synthétisés à l'aide du logiciel **Mbrola**. Nous avons effectué, dans le deuxième test, une synthèse au niveau de la durée. Autrement dit, une synthèse isochronique a été établie sur tous les stimuli naturels, c-à-d que nous avons donné la même durée pour chaque phonème afin de pouvoir vérifier l'importance du f_0 dans la distinction entre les trois attitudes. Pour cela, nous avons choisi comme durée **80 ms**. Dans le troisième test, nous avons fait l'inverse du test précédent. C'est-à-dire que la synthèse a été effectuée au niveau de f_0 . Nous avons essayé en d'autre terme d'aplatir la courbe du fondamental pour but de vérifier le rôle la durée dans la distinction entre les trois attitudes. Pour cela, une valeur de **120 HZ** a été attribuée à chaque phonème dans l'ensemble des stimuli. Il reste à signaler que la synthèse dans Mbrola a été effectuée à l'aide des scripts programmés par Cédric Gendrot.

3.2. Résultats

Comme nous l'avons signalé auparavant, les tests de perceptions étaient effectués principalement pour confirmer le rôle important du f_0 dans l'identification des attitudes en anglais américain, ainsi que montrer l'effet secondaire de la durée dans cette identification. Nous avons obtenu les résultats suivants :

- Le taux d'identification de la demande de surprise (**70%**) est le plus élevé par rapport à la demande neutre (**57%**) ou de doute (**40%**).
- La parole naturelle est la plus identifiée par rapport à la parole synthétisée. Par ailleurs, si nous comparons le taux d'identification des stimuli dans la parole

synthétisée avec une synthèse isochronique à celui de la parole naturelle, nous obtiendrons des pourcentages plus ou moins proches, ce qui confirme notre hypothèse à propos du rôle important de f_0 .

- En ce qui concerne la durée, nous constatons que l'aplatissement du fondamental a provoqué une grande perturbation chez les auditeurs, ce qui leur a mené, la plus part des auditeurs, à identifier tous les stimuli comme une demande **neutre**. Ce ci explique le taux élevé d'identification de la demande neutre. Il est important à noter que le taux d'identification des demandes de doute et de surprise, dans le troisième test, chez la majorité des auditeurs est de **0%**.

Le graphique ci-dessous résume ce qui était constaté au-dessus :

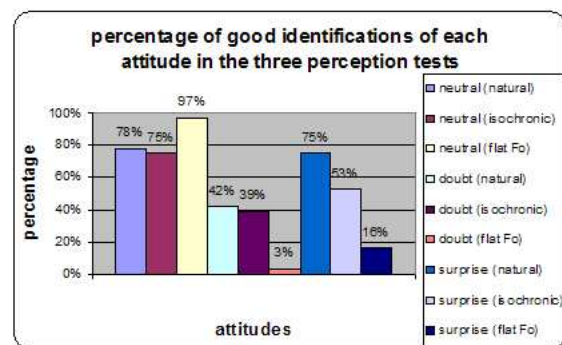


Figure 2: pourcentage de bonnes identifications de chaque attitude dans les trois tests de perceptions.

4. DISCUSSION

L'analyse acoustique de l'intonation en anglais américain a confirmé, en quelques sortes, les résultats obtenus dans d'autres études antérieures, soit en ce qui concerne la trajectoire de la ligne de base ou le rôle pivot du f_0 dans la distinction entre la parole neutre et la parole attitudinale. Par exemple, Maede [3], [4] et Vaissière [8] ont trouvé que les phrases déclaratives neutres, en français ou en anglais, suivent une trajectoire descendante qu'ils l'ont appelé la Ligne de base. Par ailleurs, nous avons conclu que cette Ligne de Base suit une trajectoire ascendante plutôt que descendante quand on ajoute une attitude à la parole neutre.

Au niveau des variations locales de f_0 , nous avons confirmé que la structure du fondamental joue un rôle très important dans la distinction entre la demande neutre, la demande de doute et la demande de surprise. Le même résultat a été obtenu par Léon [2], Morlec [6] et Yoshihori & Yoh'ichi [11] dans leurs études sur la parole neutre et attitudinale.

Ce rôle distinctif du f_0 est conditionné par plusieurs facteurs linguistiques et extralinguistiques, particulièrement l'état émotif du locuteur.

5. CONCLUSION GENERALE

En somme, cette étude a montré deux idées principales :

- Les théories de la déclinaison ne sont pas toujours conformes aux données, particulièrement celles dérivées de la parole expressive.

- La structure du fondamentale joue un rôle très important dans la distinction entre la parole neutre et la parole attitudinale.

BIBLIOGRAPHIE

- [1] E. Garding. A Generative model of intonation. *In cutler, A & Ladd, D.R.(eds). Prosody: models and measurements.* Heidelberg: Springer: 11-25, 1983.
- [2] P. Léon. Précis de phonostylistique. Parole et expressivité. Paris. Nathan, 1993.
- [3] S. Maeda. A Characterization of fundamental frequency contours of speech. *Quarterly progress report, no 114, research laboratory of electronics, M.I.T,* 1974.
- [4] S. Maeda. A Characterization of american english intonation. *PhD prepared at Massachusetts Institute of Technology,* 1976.
- [5] M. Mghizou. Intonation, question et attitude en anglais américain ; une étude acoustico-perceptive. Master's degree in phonetics. Laboratoire de Phonétique et Phonologie (UMR 7018), CNRS / Sorbonne Nouvelle, Paris, France, 2005.
- [6] Y. Morlec. Génération multiparamétrique de la prosodie du français par ap-prentissage automatique. *PhD thesis. Institut National Polytechnique de Grenoble,* 1997.
- [7] N. Thorsen. Two issues in the prosody of standard Danish. *In cutler, A & Ladd, D.R.(eds). Prosody: models and measurements.* Heidelberg: Springer: 27-38, 1983.
- [8] J. Vaissière. Contribution à la synthèse par règle du français. *Thèse de troisième cycle, université de grenoble ,* 1971.
- [9] J. Vaissière. On french prosody. *Quarterly progress report, no 114/115, 212-223 and 251-261.* Research laboratory of electronics, M.I.T, 1975.
- [10] J. Vaissière. Phonetic explanations for cross-linguistic prosodic similarities. *Phonetica* 52, 123-130, 1995.
- [11] K. Yoshinori and T. Yoh'ichi. prosodic Components of speech in the expression of emotions. *ATR Auditory and Visual Perception Research Laboratories.* The secont joint meeting of ASA and ASJ, 1988.