

Les voyelles /i/ et /y/ du français : focalisation et variations formantiques

Cédric Gendrot*, Martine Adda-Decker** et Jacqueline Vaissière*

*LPP Université Paris Sorbonne Nouvelle CNRS UMR 7018 ILPGA

** LIMSI-CNRS bât. 508, BP 133, 91403 Orsay cedex

cgendrot@univ-paris3.fr, madda@limsi.fr, jacqueline.vaissiere@univ-paris3.fr

ABSTRACT

French vowels /i/ and /y/ are known to be less subject to variation on F1/F2 plan compared to the other French vowels [7,11]. These two vowels in French are quantal and are also characterized by the proximity of two of their formants (respectively F2/F3 for /y/ and F3/F4 for /i/). Higher formants values F3 and F4 as well as their amplitude are investigated in this study in order to show that French /i/ and /y/ have more compact F4-F3 and F3-F2 respectively (and are thus more focal) as compared to seven other languages. We also suggest that these higher formants contribute to better understanding for the variation of these vowels. We underline the importance of F3 measurements to characterize the variations of these vowels, notably due to their difference in rounding.

Keywords: Quantal Theory, focalisation, vowels acoustics, formants, F3, duration, cross-languages comparison.

1. INTRODUCTION

Cette étude se place dans le cadre d'une analyse plus large des variations formantiques des voyelles du français. Dans une étude précédente [6,7], nous avons mesuré les variations de formants F1 et F2 et observé que les voyelles plus longues agrandissent l'espace acoustique F1/F2 en formant ainsi un triangle vocalique plus important comme l'indique la figure 1.

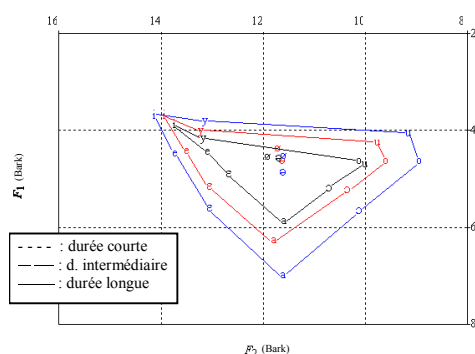


Figure 1 : Valeurs moyennes de F1 et F2 pour les voyelles orales du français en fonction de leur durée (normalisation en Bark).

Des mesures d'éloignement du centre acoustique ont révélé de manière quantitative que les variations des formants F1 et F2 sont moins larges pour les voyelles /i/ et /y/, (de même que pour les voyelles centrales /ø/ et /œ/). Nous avons alors

suggéré que ces variations étaient plus faibles pour les voyelles fermées /i/ et /y/ qui seraient, comme il est reconnu pour le français, mieux caractérisées par la proximité de F3 avec F2 (/y/) ou F4 (/i/).

Les voyelles orales du français, et parmi celles-ci les voyelles /i/ et /y/, sont fréquemment considérées, par exemple par Jones comme de bons prototypes des voyelles cardinales. /i/ et /y/ notamment sont mentionnées comme des représentantes idéales du processus de focalisation [9,10] puisque caractérisées par le rapprochement de deux de leur formants (F3/F4 et F2/F3 respectivement). Le rapprochement de deux formants a trois conséquences :

- la création d'un pic spectral proéminent dans la région fréquentielle regroupant les deux formants (l'augmentation de l'amplitude de deux formants proches est de 6 dB pour une distance deux fois plus petite entre les deux formants [4].
- les formants proches sont également intégrés d'un point de vue perceptif en un pic simple [2].
- le deuxième formant effectif (F'2 ou 'F2prime'), prenant en compte la fréquence des formants supérieurs, lors d'une synthèse des voyelles cardinales à deux formants semble correspondre à la proéminence spectrale créée par le rapprochement de ces deux formants [1].

Tabain et Perrier [11] ont également montré au moyen d'une étude analysant en parallèle les réalisations articulatoires et acoustiques du /i/ en position pré-finale de constituant prosodique que les variations du troisième formant étaient plus larges que celles des deux premiers formants, les auteurs émettant l'hypothèse que le locuteur tente d'atteindre des cibles acoustiques plutôt qu'articulatoires. L'utilisation de F'2 dans nos précédentes études [7] n'ayant pas mis en évidence des variations plus larges, nous émettons l'hypothèse que nous pourrions observer des mouvements significatifs et plus importants des formants supérieurs F3 et F4. Notons également que l'opposition entre ces deux voyelles se fait essentiellement sur la base de l'arrondissement, un arrondissement plus important qui abaisse la hauteur du troisième formant, il est ainsi probable que l'allongement des voyelles /i/ et /y/ favoriserait un rapprochement des deux formants concernés pour les voyelles les plus longues.

Le premier objectif de cette étude est de vérifier si la distance entre les formants F3 et F4 pour /i/ et F2 et F3 pour /y/ est moindre en français que dans sept autres langues étudiées précédemment (allemand, anglais américain, arabe, espagnol, italien, mandarin, portugais) [7],

ce qui pourrait être une explication de leur caractère souvent reconnu comme « cardinal ». Nous effectuons pour ce faire une comparaison avec les sept autres langues en mesurant l'écart entre F3 et F4 pour le /i/. Concernant le /y/ seulement présent en allemand et en mandarin, nous effectuerons une comparaison de l'écart entre F2 et F3.

Dans la deuxième partie de cette étude, nous vérifierons dans quelle mesure les fréquences des formants F3 et F4 sont sujettes à la variation. Cette vérification sera effectuée au moyen de la durée des voyelles mesurée sur de la parole continue. Une voyelle plus longue sera alors considérée comme plus proche de sa cible [8].

Pour finir, nous effectuerons quelques mesures sur les amplitudes des formants pour vérifier si dans le cas de voyelles plus longues, un rapprochement des formants F3/F4 pour /i/ ou F2/F3 pour /y/ génère une augmentation de l'amplitude des formants concernés. Des comparaisons seront effectuées avec les autres voyelles du français.

Pour résumer, nous tentons ici de répondre aux questions suivantes :

(i) Les voyelles focales du français /i/ et /y/ sont elles plus focales que leurs équivalents dans d'autres langues ? (ii) Cet aspect focal peut-il être confirmé pour des productions plus longues de ces voyelles grâce à un rapprochement supplémentaire des formants concernés ? (iii) Observe-t-on un réhaussement de l'amplitude des formants supérieurs pour ces voyelles ?

2. METHODE

2.1. Analyses et corpus

Lors de nos études précédentes [6,7], notre intérêt s'était porté sur les variations de formants F1 et F2. Nous avons montré qu'il était possible d'extraire de manière fiable avec des traitements automatiques un pourcentage élevé de valeurs de formants (seulement 6% de rejets dûs à des erreurs de détection). Les voyelles /i/ et /y/ étaient caractérisées par des taux plus élevés que la moyenne (7% et 15% respectivement) et jusqu'à 20% pour les /y/ les plus longs. Nos analyses se portent cette fois sur les formants supérieurs F3 et F4 que nous mesurons en suivant strictement la même procédure expérimentale que pour nos précédents travaux.

La durée mesurée des voyelles est répartie en 3 catégories (nominales) équilibrées (à savoir [30–50ms] pour les voyelles courtes, [60–80ms] pour les voyelles à durée intermédiaire et [90–110] pour les voyelles longues).

Nous effectuerons des mesures d'écarts de valeurs entre formants (F4 – F3 pour /i/, F3 – F2 pour /y/, F2 – F1 pour /a/ et /u/). Pour vérifier ces aspects, de grandes quantités de données sont nécessaires à cause de la variabilité interlocuteurs. Seuls les locuteurs masculins sont pris en compte ici puisque fréquemment plus nombreux dans les corpus analysés : une normalisation des données appliquée sur plusieurs langues étant de plus généralement déconseillée

[3]. Nous utiliserons pour le français une partie de la base de données ESTER (24 heures de parole continue) dont le détail et l'utilisation ont été mentionnés en [6], ainsi que des données utilisées précédemment [7] pour sept autres langues : allemand, anglais, arabe, espagnol, italien, mandarin, portugais (environ 2 heures analysées, soit 30000 phonèmes pour chaque langue). Tous ces enregistrements correspondent à des émissions journalistiques radio-et télédiffusées segmentées par le système d'alignement du LIMSI [5].

3. COMPARAISON INTER-LANGUES

Dans le cas de l'anglais et de l'allemand ayant une distinction entre un /i/ dit «relâché» (/i/) et un /i/ dit «tendu», seules les variantes «tendues», plus proches de celles du français, ont été considérées. Pour le mandarin, les différents tons appliqués à chaque voyelle sont fusionnés, selon la méthodologie appliquée en [7].

Table 1 : Valeurs moyennes des formants F1, F2, F3 et F4 pour /i/ incluant leurs écart-types respectifs entre parenthèses, ainsi que l'écart en Hertz entre F3 et F4.

| | F1 | F2 | F3 | F4 | F4 – F3 |
|-----------------|-----------------|-------------------|-------------------|-------------------|------------------|
| allemand | 319 (70) | 1991 (222) | 2610 (239) | 3621 (248) | 1012 (269) |
| anglais | 352 (61) | 2044 (186) | 2503 (199) | 3442 (225) | 939 (244) |
| arabe | 398 (130) | 2102 (169) | 2678 (141) | 3364 (295) | 686 (258) |
| espagnol | 375 (57) | 2126 (155) | 2784 (149) | 3634 (126) | 851 (226) |
| français | 302 (87) | 2024 (158) | 2848 (228) | 3494 (258) | 646 (230) |
| italien | 347 (61) | 2065 (231) | 2693 (236) | 3589(400) | 895 (301) |
| mandarin | 360 (109) | 2132 (358) | 2836 (290) | 3644 (265) | 809 (304) |
| portugais | 344 (67) | 1906 (185) | 2503 (277) | 3576 (277) | 1075 (329) |

Table 2 : Valeurs moyennes des formants F2 et F3 pour /y/ incluant leurs écart-types respectifs entre parenthèses, ainsi que l'écart en Hertz entre F2 et F3

| | F1 | F2 | F3 | F3 – F2 | F4 |
|-----------------|------------------|-------------------|-------------------|------------------|-------------------|
| allemand | 348 (90) | 1598 (190) | 2357 (197) | 759 (229) | 3451 (222) |
| français | 325 (124) | 1833 (154) | 2455 (211) | 622 (205) | 3271 (195) |
| mandarin | 348 (65) | 2136 (182) | 2650 (201) | 514 (214) | 3507 (145) |

Les tableaux ci-dessus indiquent que pour la voyelle /i/, l'écart F4 - F3 est le plus faible en comparaison des 7 autres langues analysées. Seul l'écart observé pour l'arabe s'approche de celui du français. Il semblerait alors trompeur d'utiliser le même symbole de l'API (par exemple le /i/ anglais, le /i/ allemand et le /i/ français) pour une même voyelle dans différentes langues puisque ils n'ont pas les mêmes caractéristiques spectrales. Le /i/ anglais par exemple est réalisé avec un schéma F1/F2 similaire à celui du français mais avec un F3 plus bas suggérant notamment une voyelle peu étirée en comparaison du français. Le /i/ anglais s'éloigne de fait des caractéristiques focales décrites en introduction. Certaines tentatives ont ainsi été effectuées [12] pour affiner la notation phonétique de ces sons, en particulier /i/ et /y/.

En ce qui concerne /y/, l'écart F3-F2 est considérablement moins important pour le français que l'allemand. Quant à la voyelle /y/ pour le mandarin, la hauteur importante (2100

Hz) du 2ème formant laisse à penser qu'il s'approche du /y/ suédois, non caractérisé par un changement d'affiliation de cavité comme c'est généralement le cas pour le français [10,12]. Cependant, le faible nombre d'occurrences recueillies (environ 150 contre plus de 1000 pour l'allemand) nous invite à considérer ces résultats avec prudence.

Dans une étude précédente [6], nous avons observé pour l'allemand que les variations des voyelles antérieures fermées telles que le /i/ et le /y/ étaient aussi larges que celles des autres voyelles, alors que l'espace acoustique est plus « fourni » dans cette zone que le français, l'allemand ayant des contreparties «relâchées» aux voyelles /i/ et /y/ notamment. Selon ces résultats, cette différence peut-être rapportée au fait que l'allemand n'est pas caractérisé par un rapprochement de F2 et F3 pour /y/ ou de F3 et F4 pour /i/ : ces voyelles ne seraient pas focales en allemand.

4. Variations en fonction de la durée de la voyelle

4.1. Variations formantiques

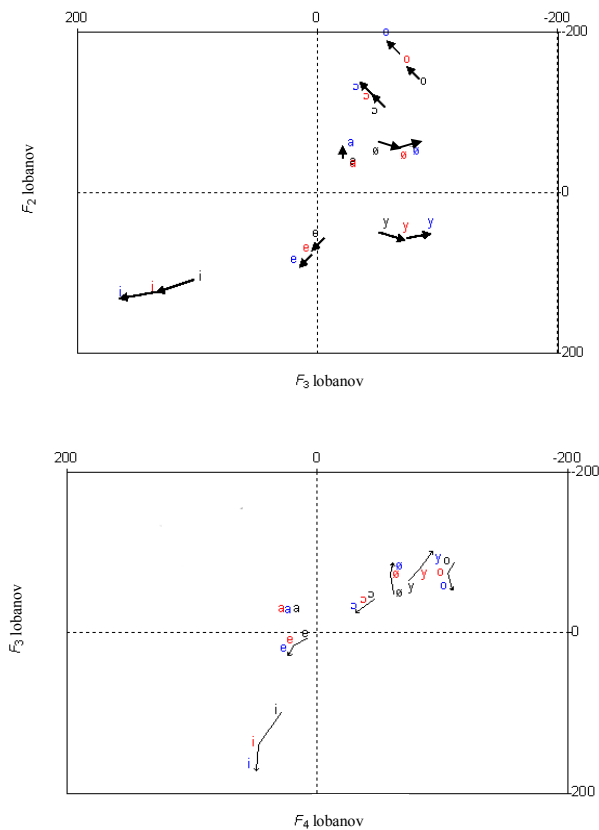


Figure 2a-b : Valeurs moyennes de F2/F3 (a) et F3/F4 (b) pour les voyelles orales du français en fonction de leur durée (normalisation Lobanov).

Nous pouvons remarquer sur la figure 2a que - lorsque la durée mesurée est plus grande - F3 augmente considérablement pour la voyelle /i/, et c'est le mouvement le plus large observé parmi toutes les voyelles.

Sur cette même figure, les variations de la voyelle /y/ sont également intéressantes puisque l'on peut observer qu'elles vont dans une direction opposée à celle des autres voyelles,

i.e. les valeurs baissent quand la durée augmente. Ces observations valent également pour /ø/ et /œ/, ce qui laisse à penser que l'arrondissement de ces voyelles, qui pousse à une baisse globale des formants, et plus spécifiquement F3, n'est pas étrangère à ce phénomène. F3 baisse lorsque la durée augmente, ce qui est contraire à ce que nous observons pour les autres voyelles. Nous l'interprétons ici comme un arrondissement supplémentaire de ces voyelles lorsqu'elles sont longues. Cela expliquerait pourquoi les mesures de dispersion depuis le centre du triangle révèlent des tendances différentes pour ces voyelles. Nous pourrions attendre dans ce cas que les autres voyelles arrondies (les voyelles postérieures) aient un 3^{ème} formant qui s'abaisse dans les mêmes conditions. Cependant les voyelles postérieures n'ont pas de contrepartie étirée en français, et leurs formants supérieurs sont caractérisés par une amplitude trop faible pour jouer un rôle perceptif.

Des résultats semblables (non développés ici) ont été trouvés pour les autres langues mentionnées dans la partie 3, à savoir une baisse importante du 3^{ème} formant lorsque la durée de la voyelle /y/ augmente, et au contraire une augmentation importante du 3^{ème} formant lorsque la durée de la voyelle /i/ augmente. Ces exemples montrent la nécessité de prendre en compte le 3^{ème} formant (au minimum) pour rendre la variabilité des voyelles (françaises) /i/ et /y/, mais également /ø/ et /œ/.

4.2. Variations d'amplitude des formants

Il serait utile de prendre un certain nombre de précautions quant aux mesures d'amplitudes de formants, notamment par des mesures relatives, puisque nous n'avons aucun renseignement sur les conditions d'enregistrement. Cependant les très faibles barres d'erreur mesurées ici laissent à penser que ces mesures sont cohérentes.

D'après les mesures brutes d'amplitudes de formants sur la figure 3, nous pouvons observer globalement pour l'ensemble des voyelles, que les amplitudes des différents formants augmentent pour les voyelles de durée intermédiaire, puis baissent pour les voyelles les plus longues. Pour la voyelle /i/ par contre, l'amplitude du 4^{ème} formant s'accroît à mesure que la durée de la voyelle augmente. Une tendance semblable peut-être observée pour l'amplitude du 3^{ème} formant, bien que moins nette. De même pour la voyelle /y/, l'amplitude du 2^{ème} formant ne décroît pas pour les durées les plus longues (mais pas pour l'amplitude du 3^{ème} formant cependant).

Ces résultats suggèrent un renforcement acoustique sensible de l'amplitude du 3^{ème} formant et du 4^{ème} formant à mesure que la durée de la voyelle /i/ augmente. Ces tendances sont moins nettes pour /y/. Des mesures relatives de différences entre les amplitudes des formants ont été effectuées pour estimer l'amplitude combinée des formants pour ces voyelles. Des mesures A1-A3 et A1-A4 indiquent des valeurs significativement plus faibles pour l'ensemble des voyelles plus longues - à l'exception des voyelles postérieures /o/ et /u/ caractérisées par une faible amplitude de F3 et F4 - suggérant ainsi une pente spectrale relevée.

Des mesures A2-A3, A2-A4 et A3-A4 indiquent des valeurs significativement plus faibles et négatives pour les /i/ plus longs et à l'inverse, plus élevées pour les /y/ plus longs. Ces résultats suggèrent une prééminence spectrale F3/F4 pour /i/ et F2/F3 pour /y/. Ces mesures devront être précisées ultérieurement : la pondération de ces valeurs d'après leur ordre et leur position dans le spectre [4] devra être effectuée afin de prendre en compte leur poids perceptif relatif.

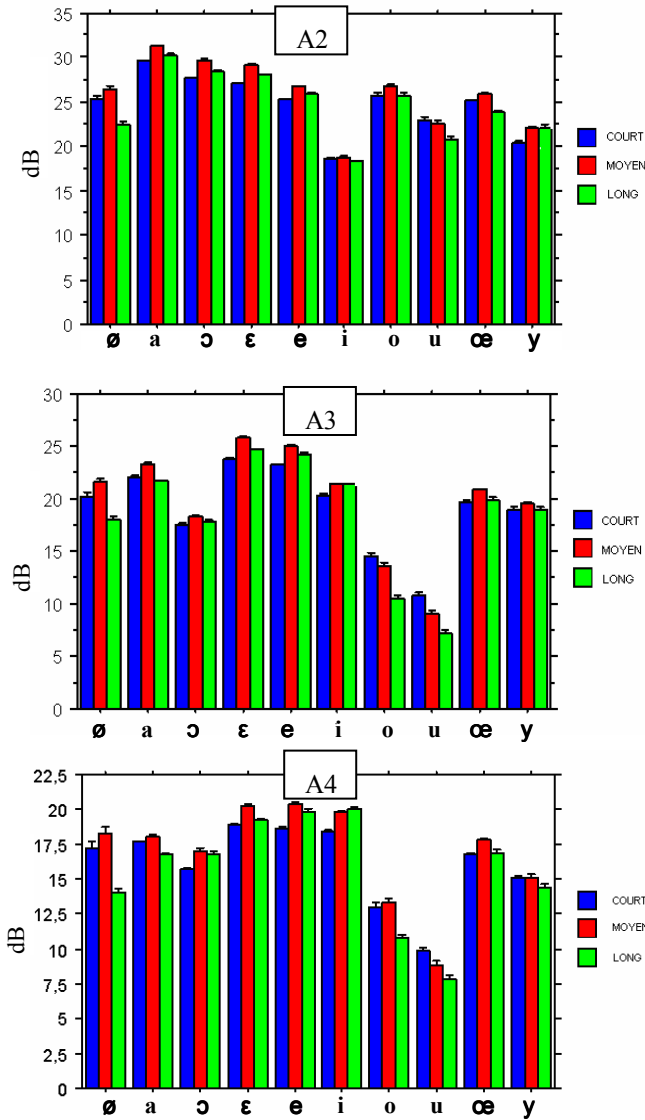


Figure 3a-b-c : Amplitude des 2^{ème}, 3^{ème} et 4^{ème} formants des voyelles orales du français en fonction de leur durée.

5. CONCLUSION

Le principal résultat de cette étude est que les voyelles /i/ et /y/ du français ont des caractéristiques focales plus prononcées que ces mêmes voyelles dans les autres langues. En effet, l'écart F4-F3 de /i/ est nettement plus faible que celui observé pour les 7 autres langues analysées. L'écart F3-F2 de /y/ est également moins large pour le français comparé à l'allemand. Par ailleurs, nous montrons ici l'importance de prendre en compte le 3^{ème} formant des

voyelles /i/ et /y/. Ce résultat est particulièrement mis en avant après avoir analysé les variations de ces voyelles en fonction de leur durée. Les variations F1/F2 mesurées précédemment comme faibles pour /i/ et /y/ en fonction de la durée de la voyelle [6] semblent ici 'compensées' par de fortes variations de F3, qui dépend des lèvres.

BIBLIOGRAPHIE

- [1] Bladon, A, Fant, G. 1978. *A two-formant model and the cardinal vowels*, STL-QPSR 1, 1-8.
- [2] Chistovich, L.A., V.V. Lublinskaja. 1979. The "center of gravity" effect in vowel spectra and critical distance between the formants: psychoacoustical study of the perception of vowel-like stimuli, *Hearing Research* 1, 185-195.
- [3] Disner, S.F. (1980) Evaluation of vowel normalization procedures. *J. Acoust. Soc. Am.* 67(1): 253-261.
- [4] Fant, G. 1960. *Acoustic Theory of Speech Production*. The Hague: Mouton.
- [5] Gauvain, J.L., Lamel, L. and Adda, G. (2002) The Limsi Broadcast News Transcription System, *Speech Communication*, 37(1-2):89-108.
- [6] Gendrot, C. et Adda-Decker, M. (2005). Impact of duration on F1/F2 formant values of oral vowels: an automatic analysis of large broadcast news corpora in French and German. *Proc. Eurospeech*, Lisbon, Portugal. pp 2453-2456.
- [7] Gendrot, C. et Adda-Decker, M. (2007). Impact of duration and vowel inventory size on formant values of oral vowels: an automated formant analysis from eight languages, *Proc. ICPHS*, Saarbrücken, Germany. pp. 1417-1420.
- [8] Lindblom, B. (1990) Explaining phonetic variation: a sketch of the H & H theory. In W. Hardcastle & A. Marchal (eds.). *Speech Production and Speech Modelling*. Dordrecht : Kluwer.
- [9] Schwartz, J.L., Boë, L.J., Vallée, N., and Abry, C. 1997. The dispersion-focalization theory of vowel systems, *J. Phonetics* 25, 255-286.
- [10] Schwartz, J.L., Beautemps, D., Abry, C., and Escudier, P. (1993). Inter-individual and cross-linguistic strategies for the production of the [i] vs [y] contrast, *J. Phonetics* 21, pp 411-425.
- [11] Tabain, M. et Perrier, P. 2005. Articulation and acoustics of /i/ in preboundary position in French. *J. Phonetics* 33-1, pp 77-100.
- [12] Vaissière, J. 2007. Area Functions and Articulatory Modeling as a Tool for Investigating the Articulatory, Acoustic, and Perceptual Properties of Sounds Across Languages. In: Sole, M-J., Beddor, P.S., Ohala, M. (eds), *Experimental Approaches to Phonology*, Oxford University Press, pp. 55-72.