

Aspects phonologique et dynamique de la distinctivité au sein des systèmes vocaliques: une étude inter-langue

Christine Meunier, Robert Espesser, Cheryl Frenck-Mestre

Laboratoire Parole et Langage – C.N.R.S. U.M.R. 6057
Université de Provence, Aix-en-Provence, France
christine.meunier@lpl.univ-aix.fr
<http://www.lpl.univ-aix.fr/~meunier/>

ABSTRACT

This study is a cross-linguistic investigation of qualitative and quantitative variations due to 1/ the structure of vocalic system, 2/ the amount of context within speech message. We hypothesize that phonetic distinctivity of vowels in a language is relative to 1/ the properties of the phonological system, 2/ the amount of informational context. Three languages (Spanish, French and English) were analyzed in three different types of speech (isolated vowels, within words and within texts). Results show 1/ centralization in the three vocalic systems relative to the amount of context, 2/ an increase of vowel dispersion also due to an increase of context information.

1. INTRODUCTION

Les systèmes vocaliques des langues du monde ainsi que la logique de leur structure ont été finement décrits (Lindblom [1], Schwartz et al. [2]). La structure de l'inventaire des voyelles dans une langue n'est pas aléatoire mais suit une logique universelle relative aux contraintes articulatoires et à la distinctivité nécessaire au sein de chaque système. Ainsi, si un système ne présente que trois voyelles, ces voyelles occuperont les espaces articulatoires les plus éloignés.

Si l'inventaire et la structuration du système (donc sa forme) suivent cette logique, on peut faire l'hypothèse que les réalisations des voyelles (donc la 'matière') vont dans le sens de cette distinctivité. Il a ainsi été montré que les locuteurs d'une langue tendent à utiliser un 'hyper-espace' de façon à rendre les voyelles très distinctes les unes des autres (Johnson [3]). Certains travaux ont montré que la densité d'un système (le nombre plus ou moins grand de voyelles) pourrait avoir une influence sur la dispersion des réalisations: plus le nombre de voyelles est réduit, plus la dispersion serait importante (Manuel & Krakow [4]). Là encore la nécessité du caractère distinct du système renforcerait cette hypothèse. Toutefois, plusieurs études ont produit des résultats inverses: dans un système à 3 voyelles (Ami), Maddieson [5] a pu montrer que les réalisations n'était pas précisément plus dispersées. De même, dans une étude récente (Meunier et al. [6]), nous avons observé une plus grande dispersion en Anglais (12 voyelles) qu'en Espagnol (5 voyelles). Nous avons alors conclu qu'il n'est pas exclu que la densité joue un rôle dans la dispersion des réalisations, mais qu'elle représente probablement un des nombreux facteurs qui conditionnent la réalisation des voyelles.

Dans certaines situations, une tendance à la réduction du système vocalique lui-même est observée. Lindblom [7] observent que les voyelles inaccentuées du suédois se centralisent lorsque le débit augmente. La persistance du phénomène de centralisation dans les systèmes vocaliques français et allemand (et non seulement spécifique aux systèmes à accent fort) a conduit Gendrot & Adda-Decker [8] à conclure qu'il s'agirait autant d'un phénomène purement physiologique (réduction de l'effort articulatoire) que d'un phénomène du à des contraintes linguistiques.

Notre hypothèse est que l'argument physiologique ne peut être une explication exhaustive du phénomène. La centralisation du système entraîne de fait une réduction de la distinctivité globale (moins de distance entre chaque élément). Cette réduction n'est possible que si d'autres aspects de la communication apporte une compensation et donc maintiennent la distinctivité à un autre niveau. Cette *variabilité* est donc *adaptive* (*H & H Theory*, Lindblom [9]). Mais plus qu'un constat d'une hypo-articulation en parole non contrôlée, il nous semble nécessaire de rendre explicite la dynamique des mouvements de variations. Nos productions sonores participent d'un équilibre entre efficacité articulatoire et préservation de la communication. Comment fonctionne cet équilibre dans le système de la langue? Notre hypothèse est qu'il est fonction d'au moins deux dimensions: l'une, que nous nommerons *Propriétés Statiques* (PS), est régie par la structure du système de sons lui-même (inventaire et propriétés du système phonologique), l'autre, nommée *Agencement Dynamique* (AD), est régie par l'ajustement ponctuel du degré d'information dans les différents secteurs de la communication (l'information présente dans d'autres secteurs linguistiques autorise une hypo-articulation). Notre objectif est de rendre explicite l'interaction entre ces deux dimensions.

2. MÉTHODE

Notre objectif étant d'évaluer les mouvements de variation des voyelles selon, d'une part, la propriété des systèmes et, d'autre part, le degré d'information contextuel, notre matériel est donc multilingue et basé sur différents types de corpus.

Trois langues sont étudiées: le français (FR), l'anglais (EN) et l'espagnol (SP). L'anglais distingue entre 13 et 15 voyelles orales alors que le français en distingue entre 10 et 12. En ce sens, la densité des deux systèmes est assez semblable, mais l'organisation du système des deux langues est assez différente. Les voyelles orales du système français se distinguent aisément avec les seuls indices F1 et F2. En

revanche, le système de l'anglais possède des indices secondaires (Vallée [10], Schwartz et al. [2]) tels que la durée et l'accent qui rendent les indices F1/F2 insuffisants pour l'identification des voyelles (Meunier et al [6]). L'espagnol comporte un inventaire comparativement moins fourni, avec seulement 5 voyelles, mais dans cet inventaire, nous trouvons les mêmes voyelles qu'en français et en anglais: /a/, /e/, /o/, /i/, /u/. Ces trois langues offrent la possibilité de distinguer l'effet de la densité (inventaire plus ou moins fourni), de l'effet de la complexité du système (indices primaires/secondaires).

Trois locuteurs (2 femmes et 1 homme) de chaque langue ont été enregistrés dans trois types de corpus différents.

- Voyelles prononcées isolément (ISO) : ce corpus permet d'évaluer la variabilité des voyelles hors contexte. Il s'agit d'une situation de production très contrôlée qui, normalement, occasionne peu de variation.
- Voyelles insérées dans des mots monosyllabiques prononcés isolément (WORD) : ce corpus permet d'évaluer la variabilité vocalique au sein du lexique. Il permet en outre d'observer des effets de coarticulation. Cette situation de production est plus naturelle et devrait occasionner une variabilité plus importante.
- Les mots du corpus WORD sont insérés dans deux textes (TXT) : ce corpus permet d'évaluer la variabilité au sein d'un contexte varié et non contrôlé. Le débit est bien sûr plus important dans ce type de corpus.

Si l'ensemble des analyses représente 3600 voyelles mesurées, le nombre de voyelles caractéristiques de chaque condition (langue, corpus, locuteur, etc.) est assez faible et limite la portée de nos résultats qui, en conséquence, nous informe essentiellement sur des tendances. Le nombre d'observations de chaque voyelle et le suivant:

- corpus ISO (5 répétitions): 10 valeurs par voyelle/langue/locuteur (il y a donc 10 mesures de /i/ pour le locuteur français LJ dans le corpus ISO)
- corpus WORD (5 rép.): 30 valeurs par voy./langue/loc.
- corpus TXT (4 rép.): 8 valeurs par voy./langue/loc.

Les analyses ont porté sur la mesure de F1 et F2 (détection automatique des formants à l'aide du logiciel ESPTS Entropics puis vérification manuelle). Deux types de variations sont observées: les variations qualitatives et les variations quantitatives (fig. 1).

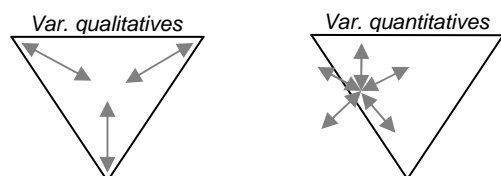


Figure 1: Type de variations observables.

Nous appelons *variations qualitatives* les variations marquées par un déplacement des cibles phonétiques. Ces variations seront observées concernant les effets de centralisation des systèmes: les valeurs moyennes se déplacent vers le centre du système. Les *variations quantitatives* caractérisent la dispersion des voyelles autour des valeurs moyennes de chaque voyelle.

L'intérêt pour nous de distinguer ces deux types de variations est d'obtenir une évaluation de la distinctivité des systèmes en fonction de différents facteurs. Par exemple, un système dont les réalisations sont centralisées et marquées par une forte dispersion des valeurs sera considéré comme peu distinctif. Cette distinctivité est évaluée en fonction de l'inventaire des systèmes (PS) et du degré de contextualisation (AD).

3. RÉSULTATS

3.1. Variations qualitatives : centralisation des systèmes

Concernant le phénomène de centralisation, nous avons porté notre attention sur les voyelles présentes dans les trois langues /i/ /e/ /a/ /o/ /u/ considérant qu'elles étaient suffisantes pour représenter fidèlement le mouvement vers le centre du système. Notons d'emblée que nos données sont hétérogènes, cela pour deux raisons principales: 1/ une importante variabilité inter-locuteur, 2/ un faible nombre d'observations pour chaque voyelle dans chaque condition. De ce fait, nous n'avons marqué significatifs (* inscrits sur les fig. 2 à 4) que les effets de variations significatifs pour chacun des trois locuteurs.

Variations qualitatives de F1 (fig. 2): on note essentiellement la fermeture des systèmes espagnols et français (F1 de /a/ diminue progressivement). Cette différence est significative pour les 3 locuteurs de chacune des deux langues (p<, 0010). Pour l'anglais, les variations ne vont pas dans le même sens pour tous les locuteurs.

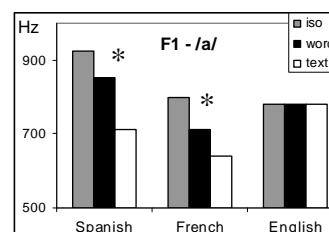


Figure 2: valeurs moyennes du F1 (en Hz) de /a/ dans les trois langues et dans les trois corpus

L'augmentation de F1 pour les voyelles fermées est une tendance qui n'est pas systématique (significative seulement pour certains locuteurs). Il est probable que cette tendance pour les voyelles fermées se confirme avec un nombre de données élargi. Notons toutefois qu'avec peu de données, la fermeture des systèmes FR et SP est bien nette.

Variations qualitatives de F2 (fig. 3-4): la centralisation de F2 est systématique pour le système espagnol aussi bien pour les voyelles antérieures (/i/ /e/) que pour les voyelles postérieures (/u/ /o/). En français, elle n'est systématique pour tous les locuteurs que pour les voyelles mi-ouvertes (diminution de F2 pour /e/, et augmentation de F2 pour /o/). En revanche, pour les voyelles fermées, la tendance est moins claire: très peu marquée pour /i/ et bien marquée pour /u/ mais seulement pour le corpus TXT. En anglais, seule l'augmentation de F2 de /o/ est significative pour tous les locuteurs. Là encore, il est probable que la tendance à la centralisation devienne plus nette avec un plus grand nombre

de données. De nouveau, le système espagnol semble plus sensible à la modalité de parole pour la centralisation.

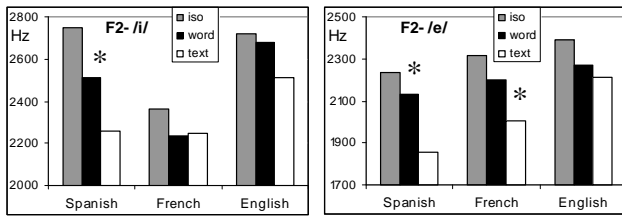


Figure 3: valeurs moyennes du F2 (en Hz) des voyelles antérieures (/i/ /e/) dans les trois langues et les trois corpus.

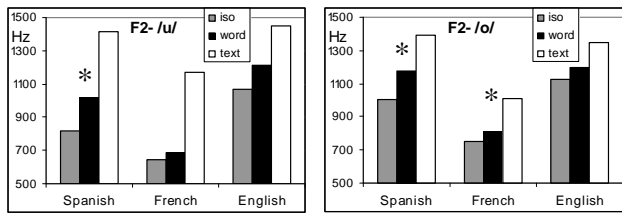


Figure 4: valeurs moyennes du F2 (en Hz) des voyelles postérieures (/u/ /o/) dans les trois langues et les trois corpus.

En résumé, on observe une tendance générale vers la centralisation des systèmes avec des effets plus marqués en fonction 1/ du système phonologique (SP>FR>EN); 2/ du type de corpus (TXT>WORD>ISO). Dans les trois langues, le système vocalique se réduit proportionnellement à la contextualisation, mais cette réduction est relative à l'inventaire phonologique de chaque langue (fig. 7 et 8).

3.2. Variations quantitatives : dispersion des voyelles

Pour obtenir une idée globale de la dispersion moyenne des voyelles dans un système, nous avons recueilli les Déviations Standards (SD) des valeurs de F1 et F2 de chaque voyelle, pour chaque locuteur et dans chaque condition. Nous avons ensuite calculé la moyenne de ces SD pour chaque langue. Ainsi, une valeur élevée sur ces figures représente une forte dispersion des voyelles.

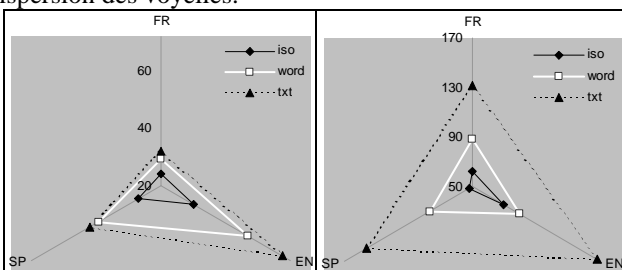


Figure 5: SD moyens du F1 (à gauche) et du F2 (à droite) des voyelles des trois langues (SP, FR et EN).

L'effet principal observé est une augmentation des valeurs des SD en fonction du degré de contextualisation (fig. 5). Plus le contexte est important (moins la situation de production est contrôlée), plus les réalisations phonétiques sont dispersées. On constate un effet grandissant de la dispersion proportionnel à la quantité de contexte: ISO>WORD>TXT. On constatera que pour les productions anglaises, le degré de dispersion est systématiquement plus grand que pour le français et l'espagnol, même si l'effet 'corpus' reste le même dans les trois langues.

3.3. Distinctivité des systèmes

Nous avons également calculé la moyenne de toutes les mesures de voyelles effectuées sur un système. Les SD de ces valeurs nous donnent une idée de la dispersion ou de la centralisation de l'ensemble du système (fig. 6). Nous observons ici que la dispersion de chaque système est plus grande en corpus ISO et plus faible en corpus TXT. Ce qui confirme nos observations sur les variations qualitatives. On remarque que, si la tendance reste la même, les systèmes espagnols et anglais évoluent différemment: les productions anglaises montrent un espace élargi en TXT qui s'accroît peu en ISO; tandis que les productions espagnoles sont très centralisées en TXT et extrêmement périphériques en ISO.

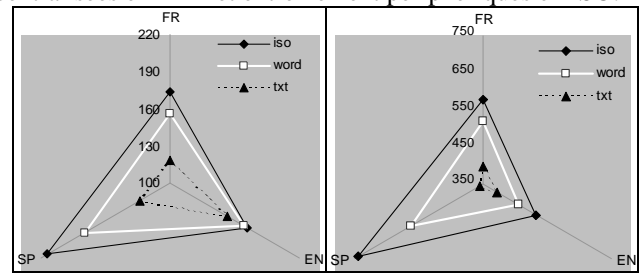


Figure 6: SD moyens du F1 (gauche) et du F2 (à droite) de l'ensemble des valeurs de chaque système.

En comparant les figures 5 et 6, on constate un effet inverse des dispersions: en TXT, dispersion maximale de chaque voyelle (fig.5) dans un espace réduit au maximum (fig. 6). Ces résultats confirment une *hyper-distinctivité* hors contexte (ISO) et une *hypo-distinctivité* en contexte (TXT). L'importance de la distinctivité est pondérée par la nature de chaque système.

4. DISCUSSION

Les résultats obtenus dans cette étude tendent à montrer que la distinctivité phonétique effective au sein des systèmes vocaliques est fonction 1/ des propriétés du système (son inventaire et le système d'indices qui le caractérise), 2/ de la quantité de contexte spécifique à la situation de production.

4.1. L'effet 'Langue'

Dans des travaux antérieurs, nous avons constaté que la densité des systèmes ne semblait pas jouer de rôle en production (Meunier et al. [6]). Ces travaux portaient sur l'observation du corpus ISO et les locuteurs espagnols montraient un hyper-espace en production et de très faibles dispersions des réalisations. Or si l'on regarde maintenant les productions en contexte, il semble que la densité joue un rôle important: le système espagnol se réduit considérablement par rapport à celui de l'anglais (fig. 7 et 8). Il semble donc que *moins un système est dense, plus il est élastique*. D'une certaine façon, le système de l'anglais est beaucoup moins sensible à la situation de production (résistance à la centralisation, dispersion déjà très importante en ISO, etc). Quand au système français il semble intermédiaire, ce qui pourrait être du à la nature de son système: aussi dense que le système anglais mais moins complexe concernant le type d'indices nécessaires à la distinction des voyelles.

4.2. L'effet 'corpus'

Au travers de cette étude, les effets des corpus sont assez nets et confirment une hyper-articulation des voyelles en corpus contrôlé (ISO) et une hypo-articulation en parole contextualisée (TXT).

L'effet systématiquement intermédiaire du corpus WORD n'est pas sans intérêt. En effet, les travaux de Lindblom [7] et de Gendrot & Adda-decker [8] tendaient à montrer que la centralisation était corrélée avec l'augmentation du débit. Nos premiers résultats concernant la durée (non exposés ici) montrent qu'il n'y a pas de différence entre les durées des voyelles ISO et celles du corpus WORD (en français, les durées WORD sont même plus longues). En revanche, la position intermédiaire de ce corpus pourrait trouver son origine dans la quantité d'information intermédiaire de l'unité lexicale.

4.3. La régulation de la distinctivité

D'une certaine façon, l'observation d'une hypo-articulation en parole non contrôlée, est juste la confirmation d'une

évidence. Il va de soi que plus la parole est contextualisée, plus le débit augmente, et plus les cibles articulatoires ne peuvent être atteintes. En revanche, il nous semble intéressant de s'interroger sur des causes et le fonctionnement de cette évidente hypo-articulation. La première cause qui nous vient à l'esprit est mécanique : la réduction vocalique serait liée à l'optimisation des contraintes gestuelles. Toutefois, cette contrainte ne peut expliquer à elle seule la diminution de la distinctivité au sein des systèmes. Cette diminution n'est possible que si l'intelligibilité du message reste intact, et donc si l'apport d'information est assuré à un autre niveau. Il semble impossible de distinguer les contraintes mécaniques des contraintes informationnelles ; elles vont de pair et l'une conditionne l'autre. Comment expliquer l'hyper-articulation en production de voyelles isolées sinon par une recherche de la distinctivité maximale dans une situation où le degré d'information apporté par le contexte est nul ? D'une certaine façon, la distinctivité du système est assurée par une hyper-articulation lorsque l'information contextuelle est absente, ou par les informations contextuelles lorsque la distinctivité articulatoire ne peut être atteinte.

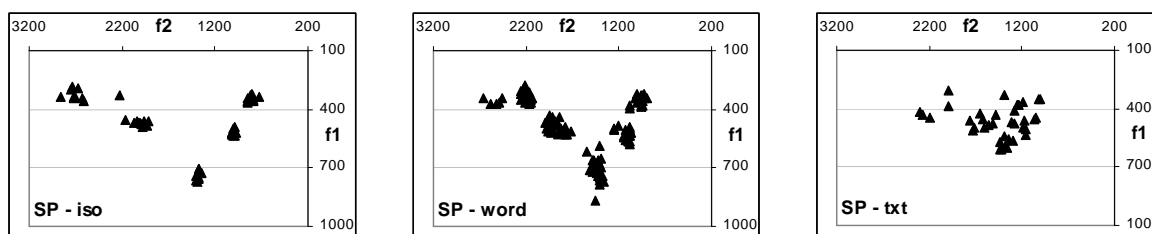


Figure 7: valeurs F1/F2 des voyelles produites par le locuteur espagnol LJ dans les trois corpus

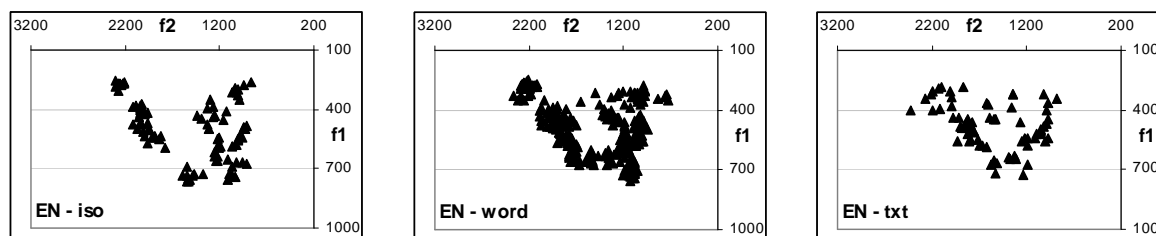


Figure 8: valeurs F1/F2 des voyelles produites par le locuteur anglais RN dans les trois corpus

BIBLIOGRAPHIE

- [1] B. Lindblom. Phonetic Universals in Vowel Systems. In *Experimental Phonology*. Edited by Ohala J.J. & Jaeger J.J., Academic Press Inc., 13-44, 1986.
- [2] J.L. Schwartz, L.J. Boë, N. Vallée. Major trends in vowel system inventories. *Journal of Phonetics*, 25, 233-253, 1997.
- [3] Johnson, K. Adaptive dispersion in vowel perception. *Phonetica*, 57, 181-188, 2000.
- [4] S.Y. Manuel, R.A. Krakow. Universal and Language particular aspects of vowel-to-vowel coarticulation. *Haskins Lab. S.R.S.R. SR-77/78*, 69-78, 1984.
- [5] I. Maddieson, R. Wright. The vowels and consonants of Amis – a preliminary phonetic report. *UCLA Working Papers in Phonetics*, Phonetics Laboratory, Los Angeles, CA, USA, 45-65, 1995.
- [6] C. Meunier, C. Frenck-Mestre, T. Lelekov-Boissard, M. Le Besnerais. Production and perception of foreign vowels: does the density of the system play a role? *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, Spain, 723-726, 2003.
- [7] B. Lindblom. Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, 35, 1773-1781, 1963.
- [8] C. Gendrot & M. Adda-Decker. Analyses formantiques de corpus radiophoniques multilingues. *Actes de la conférence MIDL 2004*, Paris, Novembre 2004.
- [9] B. Lindblom. Explaining phonetic variation: a sketch of the hyper- and hypospeech theory. *Speech Production and Speech Modelling*, Hardcastle and Marchal (eds.), Kluwer Academic Publishers, 403-439, 1990.
- [10] N. Vallée. *Systèmes vocaliques : De la typologie aux prédictions*. Thèse de Doctorat en Sciences du Langage, Université de Grenoble, 1994.