

The logo for the University of Lausanne (UNIL), featuring the word 'Unil' in a white, elegant cursive script.

UNIL | Université de Lausanne

Étude de la variation acoustique dans les voyelles du français : l'impact des conditions de production

Mélanie Lancien

Section LI, Université de Lausanne, Suisse

INTRODUCTION DU SUJET

La variation phonétique : la théorie Hyper&Hypo (H&H)

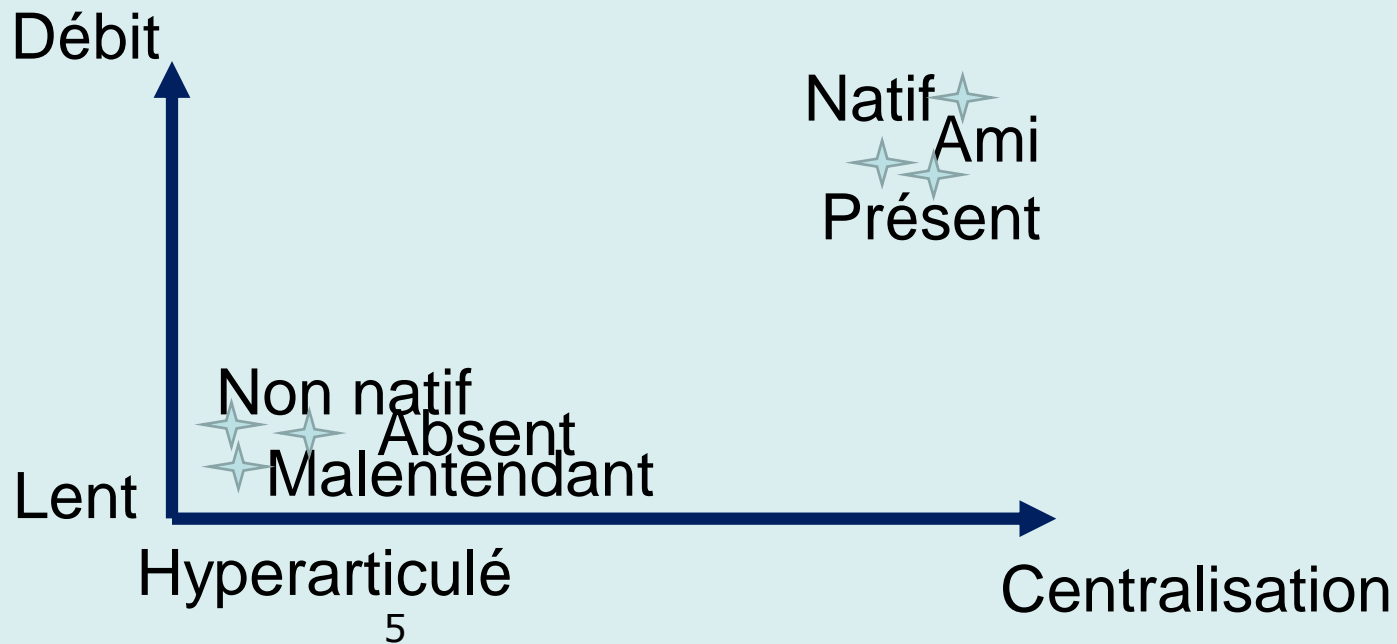
- Lindblom (1990)
 - Théorie H&H : la parole se décline sur un **continuum** qui va de l'hyper à l'hypo articulation en fonction de :
 - Besoin d'intelligibilité
 - Moindre effort articulatoire

L'influence du contexte de production

- Harmegnies & Poch-Olivé (1992, 1994)
 - Différences **spectrales** et **temporelles** entre la parole lue et la parole spontanée
 - + variations liées à la tâche/situation de production

- Scarborough et al (2007, 2013, 2015)
 - Différences acoustiques par rapport à la lecture liées à la présence / identité / langue maternelle de l'interlocuteur

Influence de l'interlocuteur



Des limites méthodologiques

- Nombre de **locuteurs** interrogés
 - Peu de locuteurs (sauf Scarborough, 2007, 2013, 2015)
- **Matériel** restreint et difficile à contrôler :
 - Quantité de données + contrôle
- Données **écologiques** :
 - Difficulté d'allier contrôle expérimental et écologie des productions

Objectifs de l'étude


Nous proposons d'étudier :


1. l'impact du style de parole sur les paramètres temporels et spectraux des voyelles orales du français, par rapport à la lecture
2. les différences entre locuteurs dans la réalisation des styles de parole

METHODOLOGIE

Protocole expérimental

- 4 binômes (8 ♀, amies, moy = 23 ans)
 - 2 sessions d'enregistrement par binôme
- 5 conditions de production :
 - Jeu de type Taboo en deux manches (indices = mots-cibles) :
 - 1^e manche : Parole libre
 - 2^e manche : 1 seul mot autorisé

Manche 1 

Manche 2 

Commente faire deviner tortue en
« terrestre » et
« verdure » ???

Joueur A
(informateur)




Tortue
Indices Bonus :
Terrestre
Verdure



Joueur B
(récepteur)




Mots-cibles

Protocole expérimental

- 4 binômes (8 ♀, amies, moy = 23 ans)
 - 2 sessions d'enregistrement par binôme
- 5 conditions de production :
 - Jeu de type Taboo en deux manches :
 - 1^e manche : Parole libre
 - 2^e manche : 1 seul mot autorisé



- Lecture simple
- Lecture rapide
- Lecture pour un mal entendant



Induit par lecture
d'un ppt

Protocole expérimental

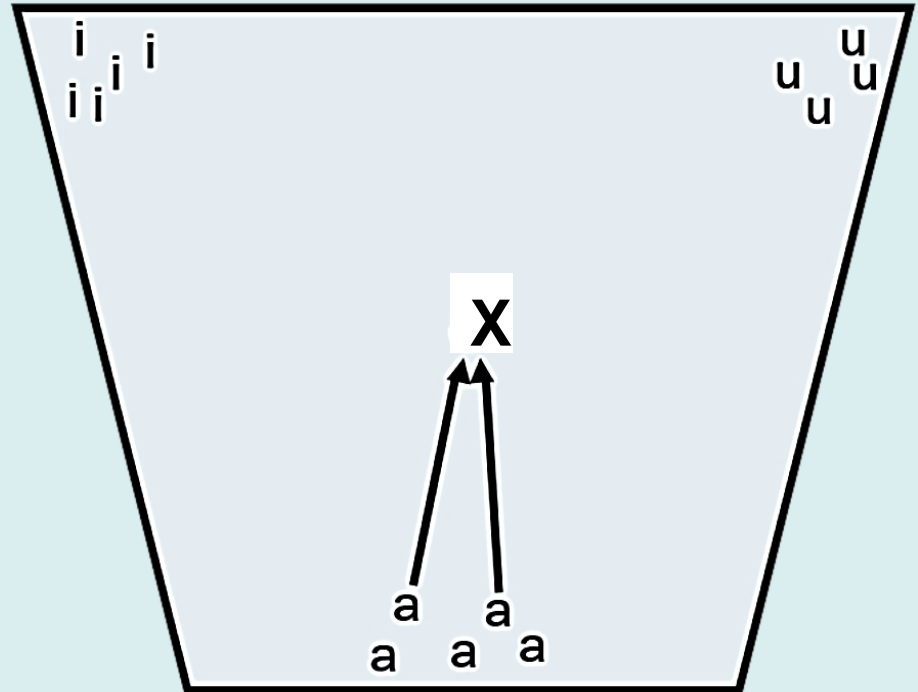
- Enregistrements :
 - Chambre sourde
 - Micro-casque

Mesures de la variation vocalique

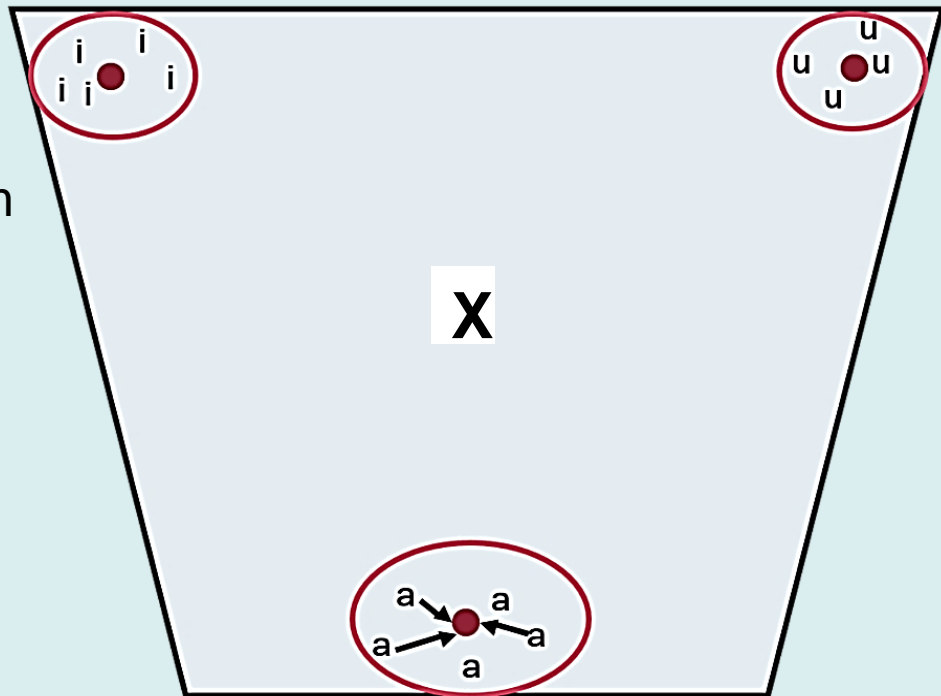
- Temporelle :
 - Durée
- **Spectrales :**

Mesures de la variation vocalique

- Temporelle :
 - Durée
- **Spectrales** :
 1. Centralisation




- Temporelle :
 - Durée
- **Spectrales** :
 1. Centralisation
 2. Dispersion



- Temporelle :
 - Durée
- Spectrales :
 1. Centralisation
 2. Dispersion
 3. Confusabilité

Catégorie prédite par le modèle

		Catégorie de la voyelle				
		a	e	i	u	o
a						
e						
i						
u						
o						


 Probabilité d'appartenance à la catégorie de chaque exemplaire (LDA)

Matériel phonétique recueilli

- 1702 exemplaires des voyelles orales /i, y, ε, œ, a, ɔ, o, u/ en seconde syllabe de mots bisyllabiques
- 22 contextes consonantiques
- Produits par 7 locutrices (~~locutrice 1~~)
- Dans 5 conditions
- 185 des 204 mot-cibles attendus produits par les 7 locutrices (soit 91%)

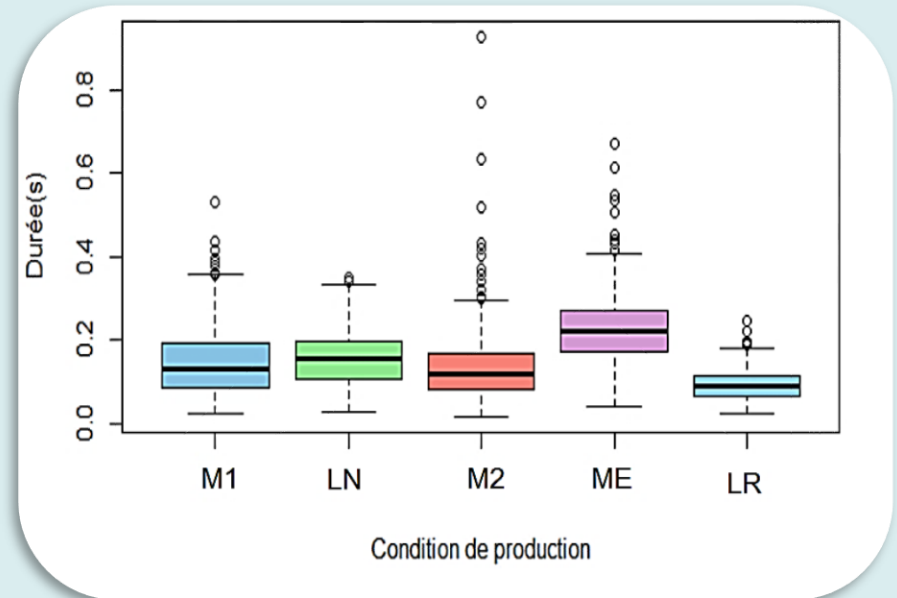
RESULTATS

L'influence de la situation de communication

Effet sur la durée

- $p < 0.001$
- $R^2 = 29\%$

- Rapide (LR) → + court
- Malentendant (ME) → + long
- Jeu : isolé (M2) = en contexte (M1)
 - isolé ≠ lecture
 - en contexte = lecture

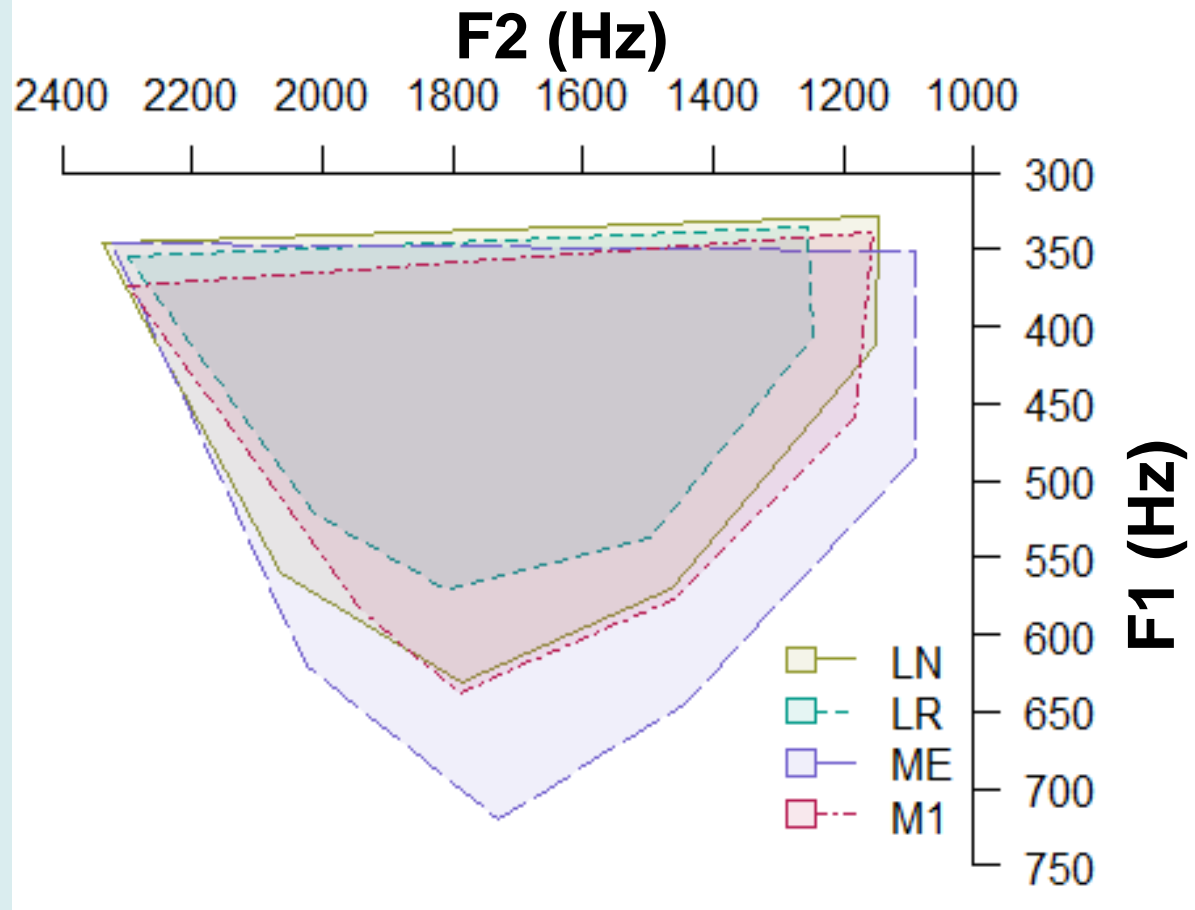


LN = Lecture Normale
LR = Lecture Rapide
ME = Lecture pour malentendant
M1 = Spont. en contexte
M2 = Spont. en isolé

*Modèle linéaire mixte, facteurs fixes = style et voyelle,
aléatoires = contexte consonantique et locuteur*

Effet sur la centralisation

- $p < 0.001$
- $R^2 = 3.65\%$
- Rapide (LR) + centralisé
- Lecture (LN) = Jeu en contexte (M1)
- Malentendant (ME) - centralisé



LN = Lecture Normale
LR = Lecture Rapide
ME = Lecture pour malentendant
M1 = Spont. en contexte
M2 = Spont. en isolé

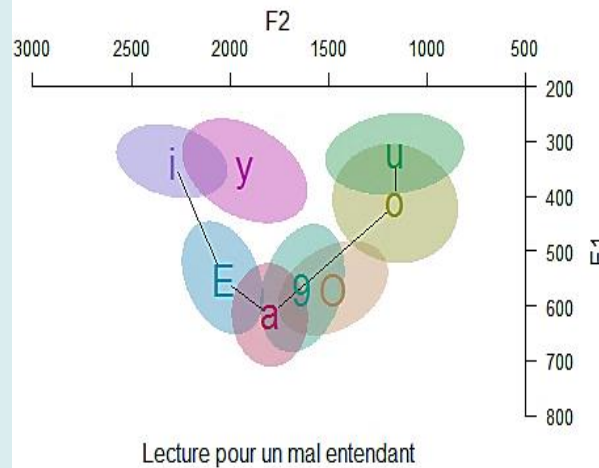
Effet sur la dispersion

- $p < 0.001$
- $R^2 = 1.48\%$

- LR=M1 : plus de dispersion
- ME=LN : moins de dispersion

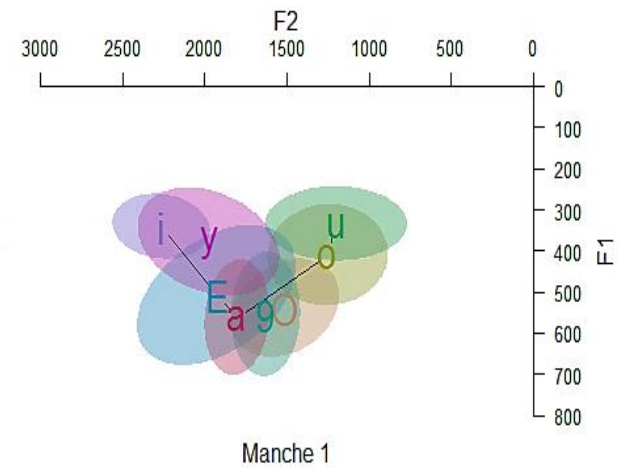
LN = Lecture Normale
LR = Lecture Rapide
ME = Lecture pour malentendant
M1 = Spont. en contexte
M2 = Spont. en isolé

LECTURE NORMALE

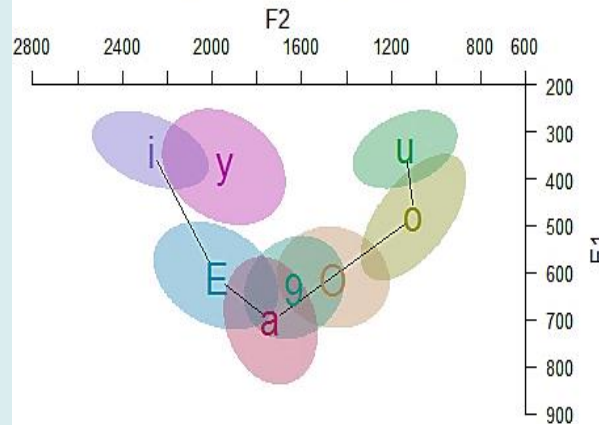


Lecture pour un mal entendant

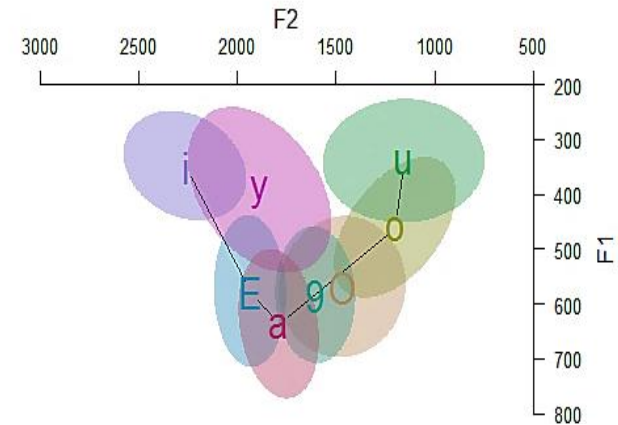
LECTURE RAPIDE



Manche 1



LECTURE ME



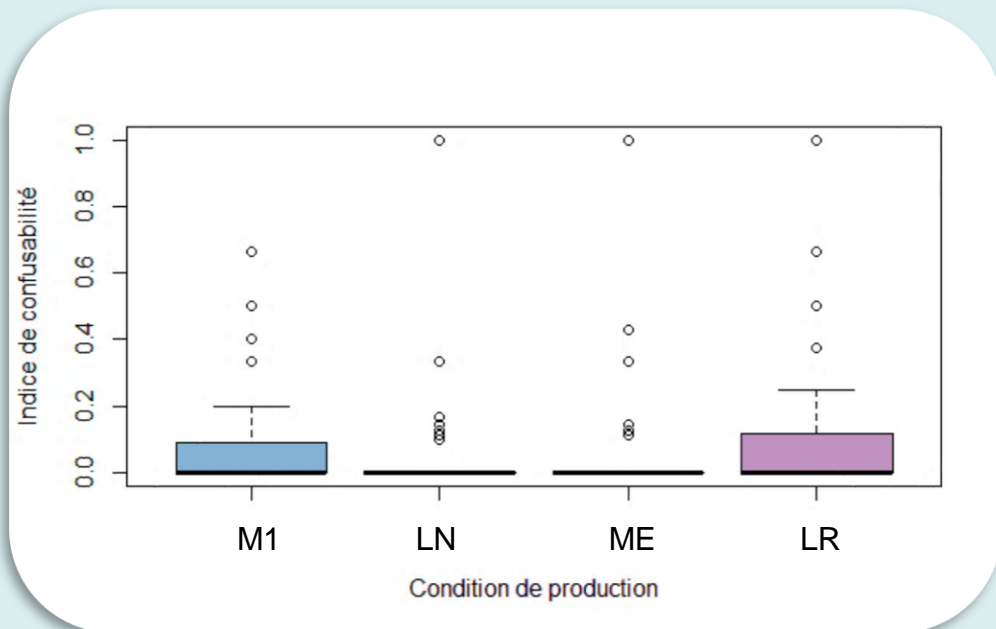
MANCHE 1

illustration en Hertz

Sur la confusabilité

pas de \neq significative entre styles

- Effet plancher (9% de mauvaise catégorisation)



LN = Lecture Normale
LR = Lecture Rapide
ME = Lecture pour malentendant
M1 = Spont. en contexte
M2 = Spont. en isolé

Corrélations entre métriques

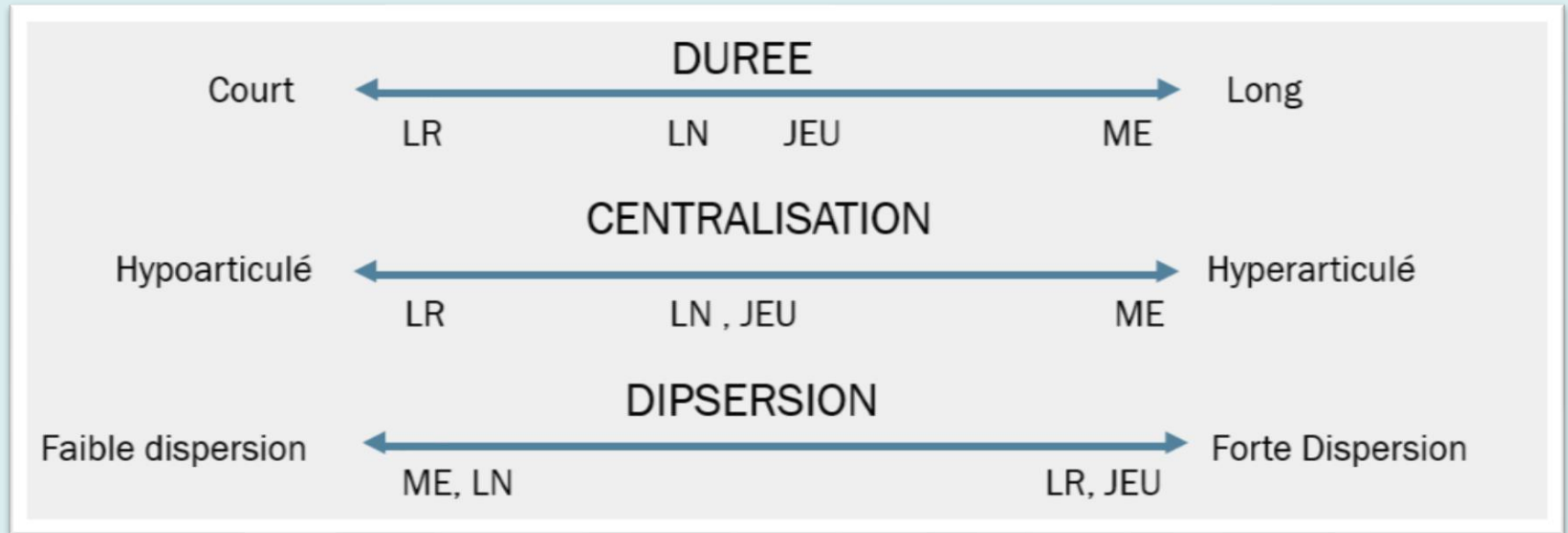
- Très faibles corrélations entre la centralisation et la dispersion ($r < 0.3$)
 - Centralisation pas toujours accompagnée d'une dispersion plus faible
- Très faible corrélation entre la durée et les mesures spectrales à l'intérieur de chaque style ($r < 0.25$)
- Les métriques sont corrélées différemment pour chaque style

L'effet du locuteur

- Effet significatif du locuteur ($p < 0.001$)
 - Forte variation inter-locuteur
 - Des différences de réalisation des styles surtout entre types de lecture
 - **Jeu (mots-cibles en contexte) → peu de différences inter-locuteurs**
 - Effet principal du locuteur uniquement sur la durée et la centralisation

DISCUSSION ET CONCLUSIONS

Positionnement des styles sur le continuum hyper-hypo

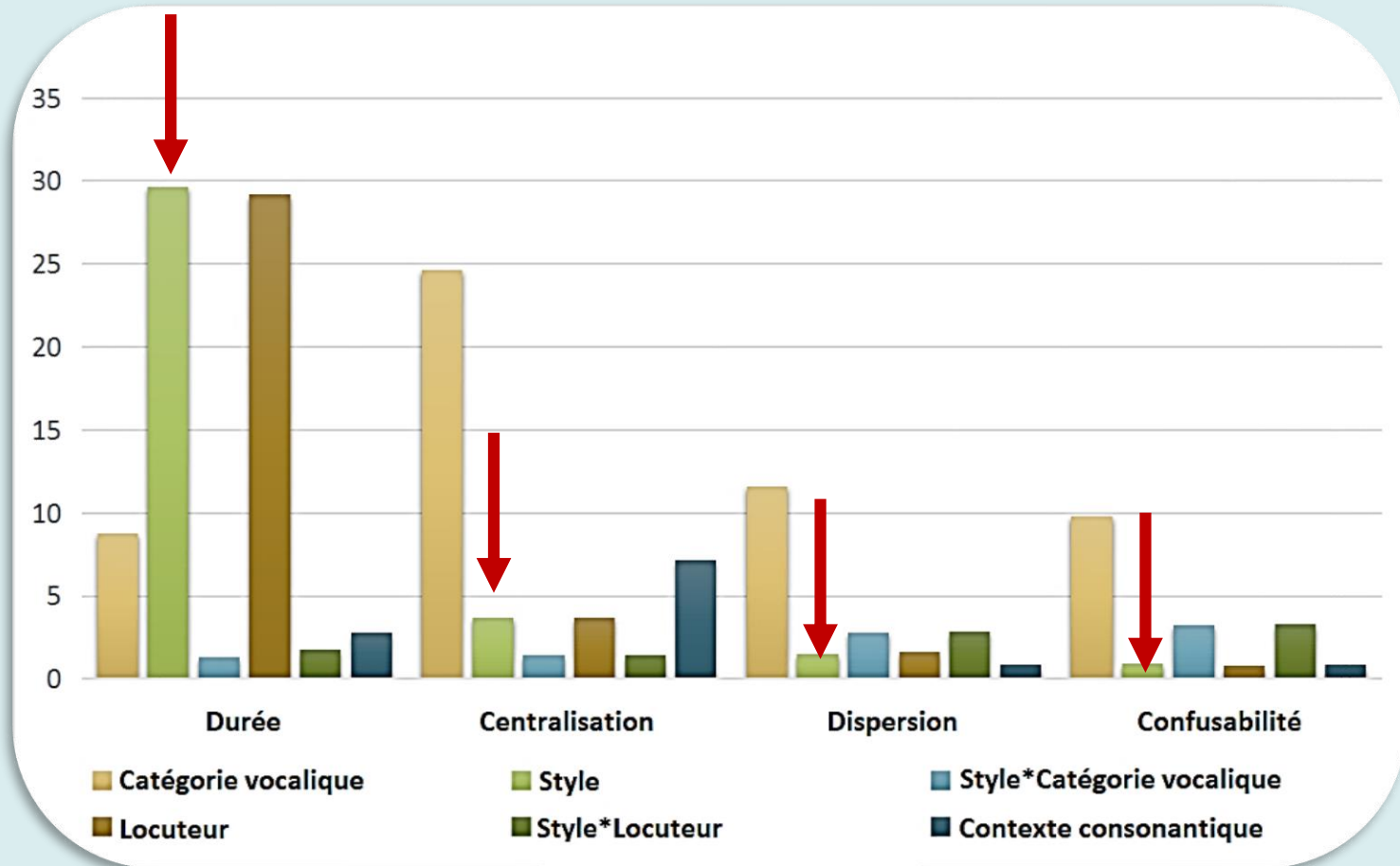


Effet du style sur la confusabilité mais
pas de différences significatives entre styles

LN = Lecture Normale
LR = Lecture Rapide
ME = Lecture pour malentendant
M1 = Spont. en contexte
M2 = Spont. en isolé

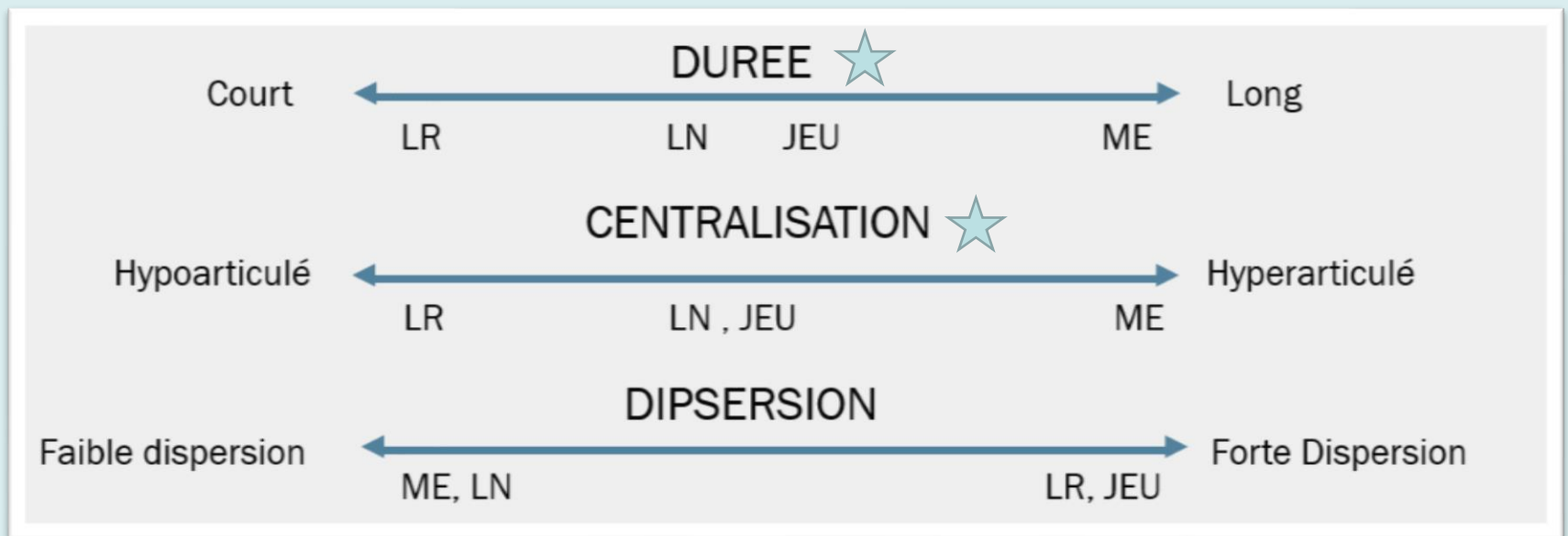
Tailles d'effet

Valeurs R^2 d'explication de la variation (Nakagawa et Cuthill, 2007)



Comparaisons avec la littérature

★ = Scarborough et al (2007, 2013)



LN = Lecture Normale
LR = Lecture Rapide
ME = Lecture pour malentendant
M1 = Spont. en contexte
M2 = Spont. en isolé

≠ Mesures de Bradlow (1996) sur la réduction formantique :
expansion plus présente sur l'axe F2 que sur l'axe F1 en
ME

Nous n'avons pas retrouvé ce phénomène



Multidimensionnalité de la variation vocalique

- Les relations entre métriques montrent la multidimensionnalité de la réduction spectrale et temporelle
 - Jeu (contexte) : dispersion intra-catégorie + et centralisation +
 - Faibles corrélations entre les dimensions de la variation vocalique

Conclusions Générales

- Répartition différente des styles sur l'axe hyper-hypo
- Multidimensionnalité de la réduction vocalique
- Stratégies différentes selon les locuteurs

Prolongement

- Variation dans la parole
 - Informations **paralinguistiques** portées par le signal (affects produits en interaction + informations indexicales)
 - **Observable à travers les métriques caractéristiques de l'hyper/hypo-articulation**
- Différents **phonostyles** (Léon 1993, Fonagy 1983, ...)
 - Lecture
 - Discussion guidée
 - Discussion libre
 - + prise en compte de l'identité de l'interlocuteur
- Approche expérimentale
- Corpus PFC - Canada

/ mɛʁsi /

Bibliographie

- Bradlow, Ann R et Torretta, G. M. e. P. D. B. (1996). Intelligibility of normal speech i : Global and fine-grained acoustic-phonetic talker characteristics. *Speech communication*, 20(3) :255–272
- Blumstein, S. E., & Stevens, K. N. (1979). Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *The Journal of the Acoustical Society of America*, 66(4), 1001-1017.
- Carlson, Rolf et Gustafson, K. e. S. E. (2006a). Cues for hesitation in speech synthesis. In Ninth Inter- national Conference on Spoken Language Processing.
- Carlson, Rolf et Gustafson, K. e. S. E. (2006b). Modelling hesitation for synthesis of spontaneous speech. *Proc. Speech Prosody 2006*.
- Erickson, Donna et Abramson, A. e. M. K. e. K. T. (2000). Articulatory characteristics of emotional utterances in spoken english. *Adm*, 98 :1–6.
- Erickson, Donna et Fujino, A. e. M. T. e. M. C. e. Y. K. e. S. Y. (2004a). Articulation of sad speech : comparison of american english et japanese. In *Acoust. Soc. Japan, Fall Meeting*.

- Erickson, Donna et Menezes, C. e. F. A. (2004b). Some articulatory measurements of real sadness. In Eighth International Conference on Spoken Language Processing.
- Erickson, D. (2006a). Some gender and cultural differences in perception of affective expressions. In Third International Conference, Dresden, Germany.
- Erickson, Donna et Yoshida, K. e. M. C. e. F. A. e. M. T. e. S. Y. (2006b). Exploratory study of some acoustic and articulatory characteristics of sad speech. *Phonetica*, 63(1) :1–25.
- Fant, G. (1960). *Acoustic theory of speech production: with calculations based on X-ray studies of Russian articulations* (Vol. 2). Walter de Gruyter.
- Fónagy, I. (1983). *La vive voix : essais de psycho-phonétique*, volume 20. Payot.
- Georgeton, Laurianne et Paillereau, N. e. L. S. e. G. J. e. K. T. (2012). Analyse formantique des voyelles orales du français en contexte isolé : à la recherche d'une référence pour les apprenants de fle. In Conférence conjointe JEP-TALN-RECITAL 2012, pages 145–152.
- Harmegnies, B et Poch-Olivé, D. (1994). Formants frequencies variability in french vowels under the effect of various speaking styles. *Le Journal de Physique IV*, 4(C5) :C5–509.
- Labov, W. (1972). *Sociolinguistic patterns*. Number 4. University of Pennsylvania Press.
- Léon, P. R. (1993). *Précis de phonostylistique : parole et expressivité*. Nathan.
- Lindblom, B. (1990). Explaining phonetic variation : A sketch of the h&h theory. In *Speech production et speech modelling*, pages 403–439. Springer.

- Passy, P. (1917). Les sons du français.
- Poch-Olivé, D et Harmegnies, B. (1992). Variations structurelles des systèmes vocaliques en français et espagnol sous l'effet du style de parole. *Le Journal de Physique IV*, 2(C1) :C1-283.
- Scarborough, Rebecca et Zellou, G. (2013a). Clarity in communication : "clear" speech authenticity and lexical neighborhood density effects in speech production and perception. *The Journal of the Acoustical Society of America*, 134(5) :3793-3807.
- Scarborough, Rebecca et Dmitrieva, O. e. H.-L. L. e. Z. Y. e. B. J. (2007). An acoustic study of real and imagined foreigner-directed speech. *Journal of the Acoustical Society of America*, 121(5) :3044.
- Scarborough, R. (2013b). Neighborhood-conditioned patterns in phonetic detail : Relating coarticulation and hyperarticulation. *Journal of Phonetics*, 41(6) :491-508.
- Stevens, K. N. (2000). *Acoustic phonetics* (Vol. 30). MIT press.
- Thibeault, M. (2011). Les émotions : une étude articulatoire, acoustique et perceptive. PhD thesis.
- Tubach, F. C. (1969). *Index exemplorum: A handbook of medieval religious tales* (No. 204-205). Suomalainen tiedeakatemia.
- Williams, Carl E et Stevens, K. N. (1972). Emotions et speech : Some acoustical correlates. *The Journal of the Acoustical Society of America*, 52(4B) :1238-1250.
- Zellou, G., & Scarborough, R. (2015). Lexically conditioned phonetic variation in motherese: age-of-acquisition and other word-specific factors in infant-and adult-directed speech. *Laboratory Phonology*, 6(3-4), 305-336.