

# Représentations acoustiques de la parole

Nicolas Audibert

Laboratoire de Phonétique et Phonologie, Paris

UMR7018 CNRS & Sorbonne Nouvelle



*Conférence invitée*

*Rencontres Jeunes Chercheurs en Parole, Grenoble, 29 novembre 2023*

# Comment mesurer/quantifier/qualifier le signal acoustique de parole ?

- Question récurrente en phonétique et en sciences de la parole
- Approche dépendante du domaine d'application (schématiquement phonétique / traitement automatique, mais porosité croissante)
- Au-delà des paramétrisations globales pour l'apprentissage statistique, mesures dépendantes du type d'unité (segment/syllabe/UIP, etc.) analysé et des phénomènes à caractériser

# Comment mesurer/quantifier/qualifier le signal acoustique de parole ?

- Dans la grande majorité des cas, mesures qui consistent en une projection de l'information spectrale (et/ou de l'analyse de la périodicité), souvent guidée par la segmentation et mise en lien avec la durée
- Par exemple :
  - Formants pour les voyelles
  - Moments spectraux pour les fricatives
  - VOT pour les occlusives
  - Proportion de la durée totale détectée comme voisée (V-ratio) pour les consonnes sonores

# Quelques exemples de caractérisation acoustique au niveau segmental

- Voyelles
  - distinction /e/ vs. /ɛ/ en position finale
  - espace vocalique et styles de parole
  - harmonie vocalique
  - effet de l'âge sur l'espace vocalique
  - (parole et chant dirigé vers l'enfant)
- Consonnes
  - réalisation comme approximantes des fricatives sonores

# VOYELLES

# Espace vocalique et styles de parole

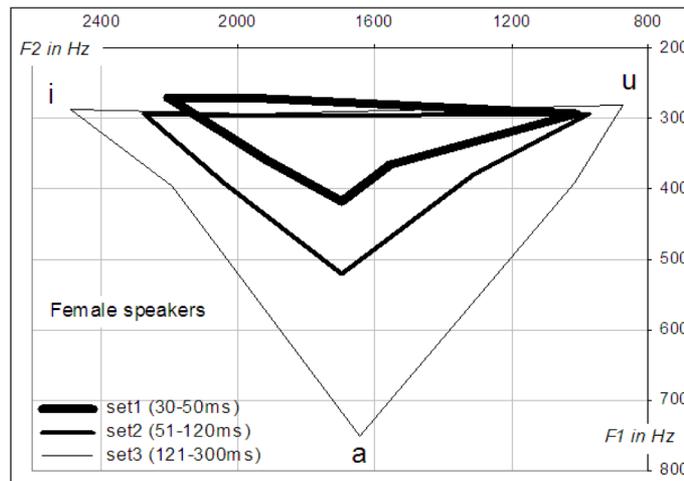
Audibert, N., Fougeron, C., Gendrot, C., & Adda-Decker, M. (2015). *Duration-vs. Style-Dependent Vowel Variation: a Multiparametric Investigation*. Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS'15). Actes en ligne (paper number 0753).

# Principaux facteurs de variation phonétique au niveau segmental

- Deux facteurs interdépendants largement décrits
  - Durée segmentale
  - Style de parole / situation de communication
- “clear speech” ↔ Parole conversationnelle
  - hyperarticulation* ↔ *hypoarticulation*
  - segments longs* ↔ *segments courts*
  - espace vocalique large* ↔ *espace vocalique réduit*
- Lien durée ↔ précision de l’articulation  
considéré comme un compromis biomécanique  
(Lindblom, 1990; Moon & Lindblom, 1994)

# Réduction vocalique dépendante de la durée en français

- Pour un même style de parole, plus de réduction sur les voyelles brèves



Parole conversationnelle  
(Meunier & Espesser, 2012)

- courtes
- moyennes
- longues

- Résultats comparables pour la parole journalistique  
(Gendrot & Adda-Decker, 2005)

# Une variation vocalique dépendante du style de parole ?

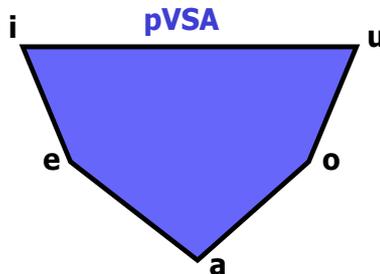
- Résultats qui suggèrent une réduction de l'espace vocalique dépendante du style de parole
    - Centralisation globale en parole conversationnelle (Ferguson & Kewley-Port, 2007)
    - Réduction qui touche principalement le F1 des /a/ en parole journalistique (Gendrot & Adda-Decker 2005)
    - Différences entre styles de valeurs MFCC uniquement sur les segments courts (Rouas, Beppu & Adda-Decker, 2010)
- ⇒ Objectif de l'étude présentée : comparer les influences respectives de la durée et du style de parole sur la variation de l'espace vocalique

# Comment quantifier la réduction vocalique ?

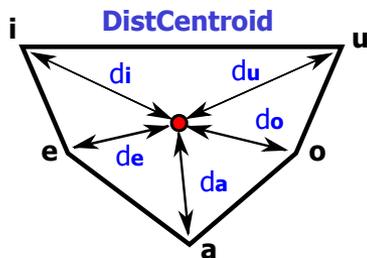
- *Reduction or expansion of a vowel space is "neither uniform nor simple"* (Ferguson & Kewley-Port, 2002, 2007, voir aussi Harmegnies & Poch-Olivé, 1992)
- ⇒ Multiples dimensions relatives aux valeurs formantiques à prendre en compte
- Centralisation / réduction de l'espace vocalique
  - Dispersion au sein de chaque catégorie de voyelle
  - Neutralisation des contrastes entre catégories
- Dimensions corrélées, mais seulement jusqu'à un certain point...*

# Mesure des variations vocaliques : Centralisation / réduction

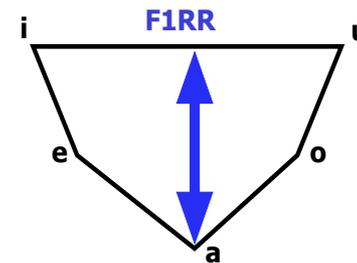
- Points de référence : centroïdes des catégories
- Générale
  - Spécifique à une dimension acoustique (Sapir et al. 2010)



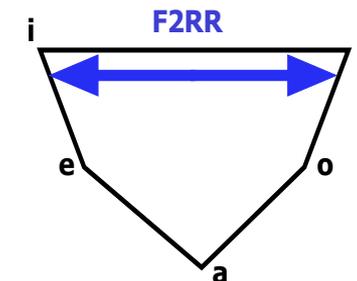
*Aire du polygone*



*Distance moyenne  
au centroïde*

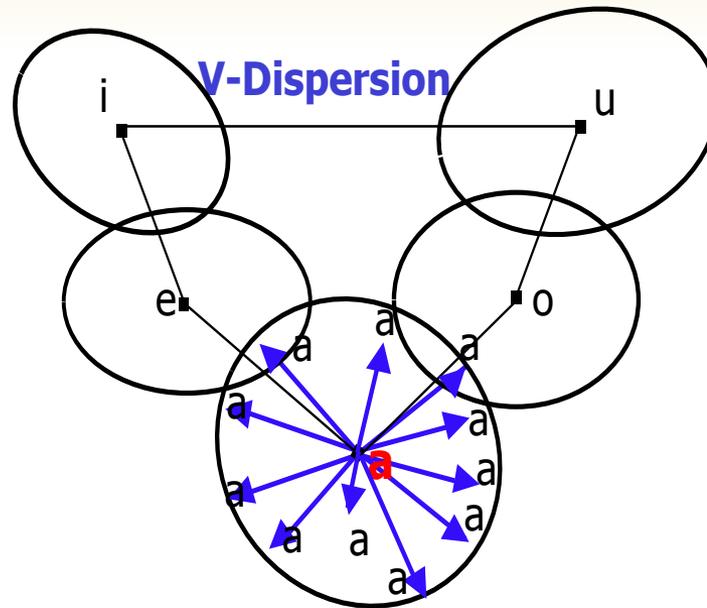


*Amplitude sur F1*



*Amplitude sur F2*

# Mesure des variations vocaliques : Dispersion intra-catégories

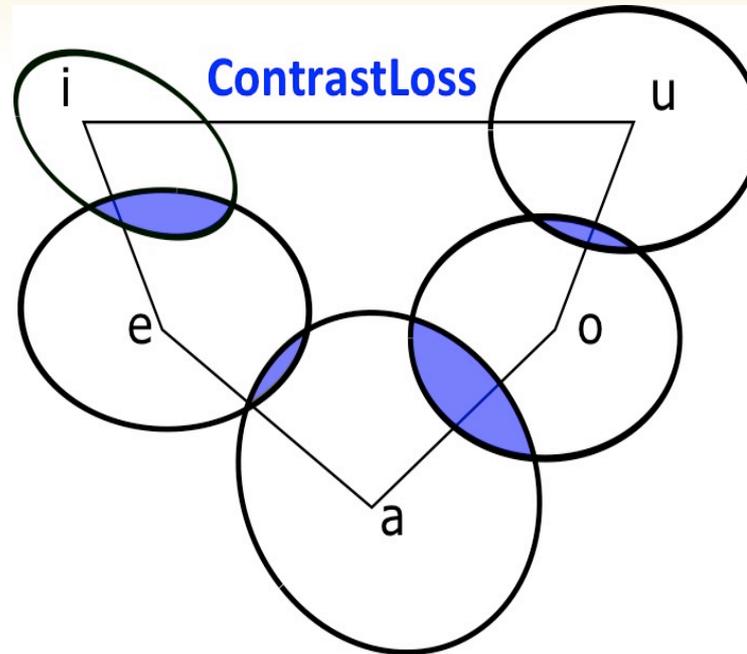


*Aire moyenne des ellipses  
de confiance à 95%*

*(ellipse qui inclut 95% des valeurs en considérant une distribution normale bivariée)*

*Alternative : distance moyenne au centroïde de chaque catégorie*

# Mesure des variations vocaliques : Neutralisation des contrastes



*Aire moyenne des intersections entre ellipses,  
estimées par échantillonnage*

*Non retenue car moins distinctive  
que la métrique suivante*

# Mesure des variations vocaliques : Neutralisation des contrastes

catégorie effective

|   | a | e | i | u | o |
|---|---|---|---|---|---|
| a |   |   |   |   |   |
| e |   |   |   |   |   |
| i |   |   |   |   |   |
| u |   |   |   |   |   |
| o |   |   |   |   |   |

prédit par F1 et F2

*% de mauvaise classification à partir de F1 et F2  
à partir d'une analyse linéaire discriminante (LDA)  
(d'après Harmegnies & Poch-Olivé, 1992)*

# Mesure des variations vocaliques : Neutralisation des contrastes

catégorie effective

|   | a | e | i | u | o |
|---|---|---|---|---|---|
| a |   |   |   |   |   |
| e |   |   |   |   |   |
| i |   |   |   |   |   |
| u |   |   |   |   |   |
| o |   |   |   |   |   |

*% mauvaise classification à partir de F1 et F2  
à partir d'une analyse linéaire discriminante (LDA)  
(d'après Harmegnies & Poch-Olivé, 1992)*

# Calculateur de métriques



## Exemplar-wise vowel space metrics computation

Nicolas Audibert

Laboratoire de Phonétique et Phonologie, UMR7018 CNRS/Sorbonne Nouvelle

R implementation made available as a Shiny app of the vowel-space metrics DistCentroid, VDispersion and ContrastLoss used in the following articles:

- Audibert, N., Fougeron, C., Gendrot, C., & Adda-Decker, M. (2015). *Duration-vs. style-dependent vowel variation: A multiparametric investigation*. In *18th International Congress of Phonetic Sciences (ICPhS 2015), Glasgow, United Kingdom*
- Hermes, A., Audibert, N., & Bourbon, A. (2023, accepted). *Age-related vowel variation in French*. In *20th International Congress of Phonetic Sciences (ICPhS 2023), Prague, Czech Republic*

Due to limited server capacity, the size of the processed data files cannot exceed 10MB.

The R code of the non-interactive version used in the ICPhS 2023 paper can be downloaded here: [download link](#)

### Select input file

acoustic\_analysis\_aging\_ICPh

### Group column:

spkCode ▼

### Vowel column:

label ▼

### Formants columns selection regex

F[1-3]

### Regex type

regex ▼

Select

| ColNum | ColLetter | Colname          | Selected                 |
|--------|-----------|------------------|--------------------------|
| 1      | A         | textgrid_file    | <input type="checkbox"/> |
| 3      | C         | spkCodeRemainder | <input type="checkbox"/> |
| 4      | D         | firstName        | <input type="checkbox"/> |
| 5      | E         | uttLength        | <input type="checkbox"/> |

Convert values from Hertz to Bark

### Distance type:

euclidean ▼

Include centroid values in exported data

Include LDA predicted category with probability in exported data

# Corpus et méthode :

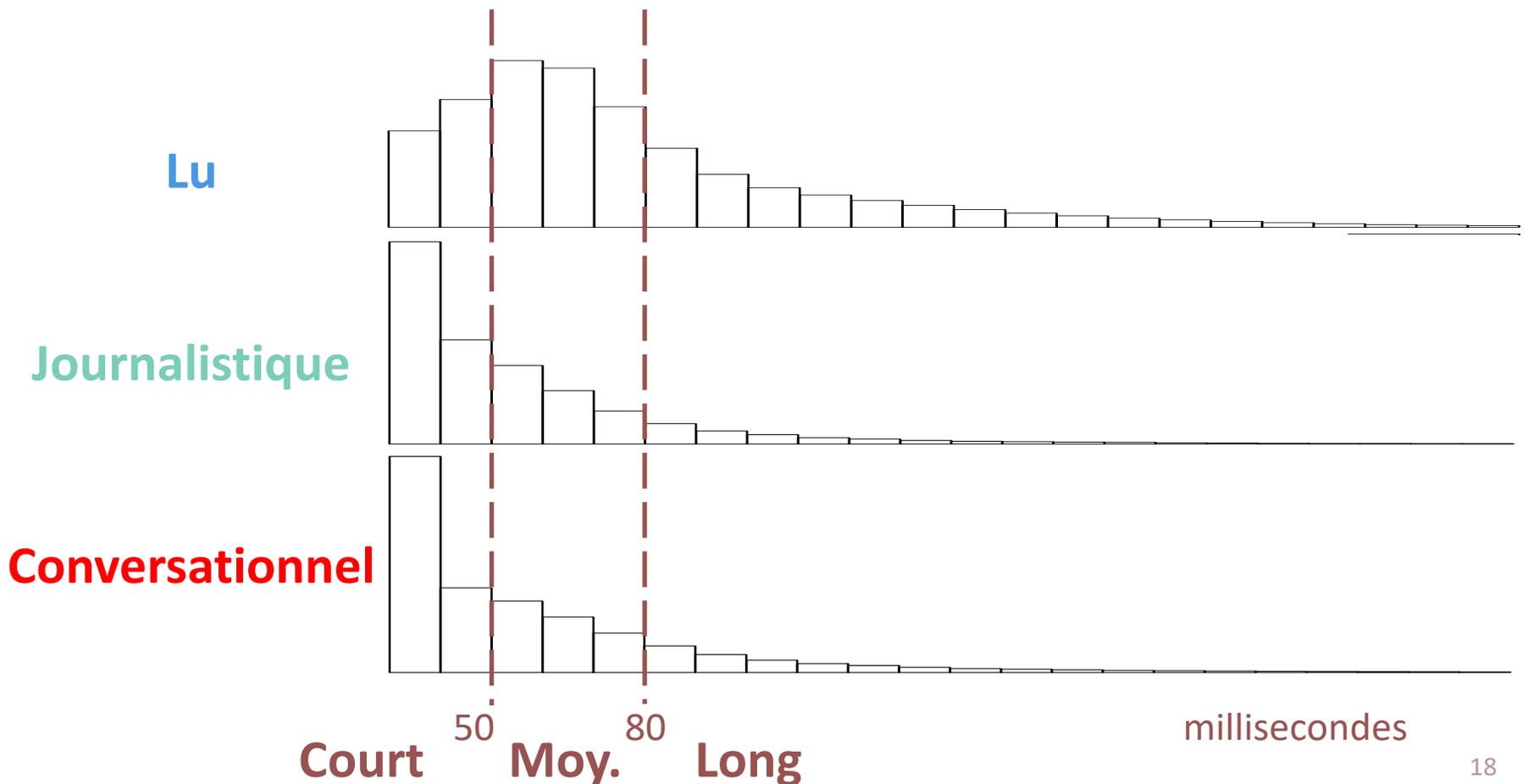
## le facteur STYLE

- Analyse et comparaison de 1 144k voyelles extraites de 3 grands corpus en français

| Corpus   | Style  | # loc | # voy |
|--|--|-------|-------|
|  <b>BREF</b><br>(Lamel et al. 1991)        | <b>Lu</b><br>phrases du journal <i>Le Monde</i> , locuteurs non-professionnels                                   | 74    | 592k  |
| <b>ESTER</b><br>(Galliano et al., 2009)  | <b>Journalistique</b><br>parole préparée, locuteurs professionnels : informations + débats politiques/de société | 61    | 233k  |
|  <b>NCCFr</b><br>(Torreira et al., 2010) | <b>Conversationalnel</b><br>conversations face à face (libres + guidées) entre amis                              | 45    | 319k  |

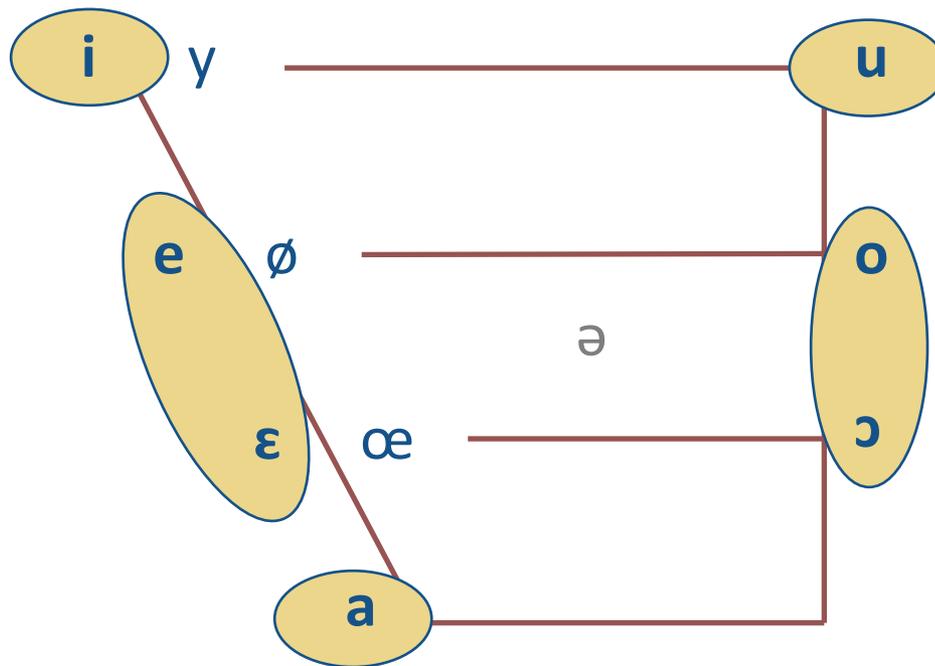
# Corpus et méthodes : facteur DUREE

- 3 classes de durée dans chaque corpus :



# Sous-ensemble des voyelles orales du français

- 5 catégories de voyelles utilisées pour les calculs
- Mi-ouvertes et mi-fermées fusionnées du fait de leur neutralisation fréquente en français



# Corpus et méthodes : valeurs formantiques

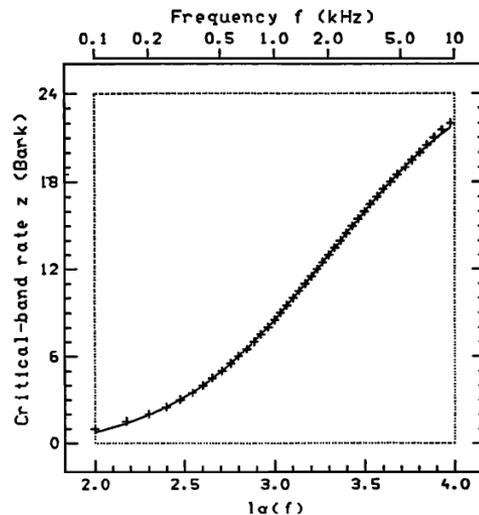
- Frontières temporelles des voyelles obtenues par un alignement forcé
  - Transcription orthographique
  - Phonétisation auto. (dictionnaire + règles)
  - Association des symboles successifs aux frontières des segments correspondants à l'aide de modèles acoustiques des sons et de leur contexte
- Alignement sensible notamment à
  - Fidélité de la transcription et de la phonétisation (ex. : reprises, hésitations, formes réduites absentes du dictionnaire phonétisé)
  - Correspondance entre les exemplaires représentés dans les modèles acoustiques et les signaux à aligner

# Corpus et méthodes : valeurs formantiques

- Valeurs F1 et F2 extraites par Praat au milieu de chaque voyelle  
(méthode de Burg)
- Filtrage des valeurs formantiques aberrantes (crible défini d'après Gendrot & Adda-Decker, 2005)  
Causes possibles des erreurs :
  - Décalage entre symboles/frontières attribuées par l'alignement automatique et signal
  - Erreurs de l'algorithme de détection des formants (par exemple F3 confondu avec F2)=> Elimination de 3% à 7% des voyelles

# Corpus et méthodes : valeurs formantiques

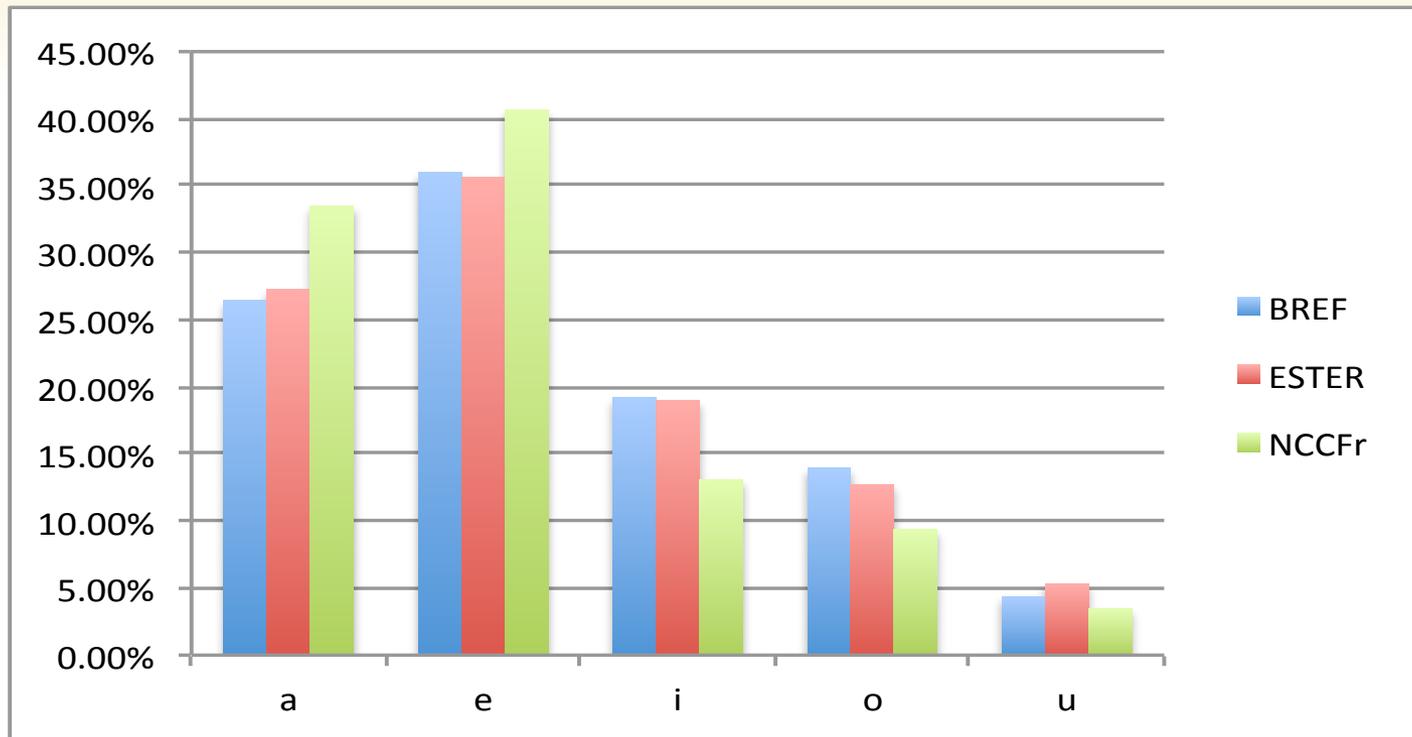
- Conversion des fréquences formantiques de Hertz en Bark (Traunmüller, 1990)



Tables de correspondances issues  
d'expériences de psychoacoustique,  
approximées par l'équation :  
$$z = [26.81/(1 + 1960/f)] - 0.53$$

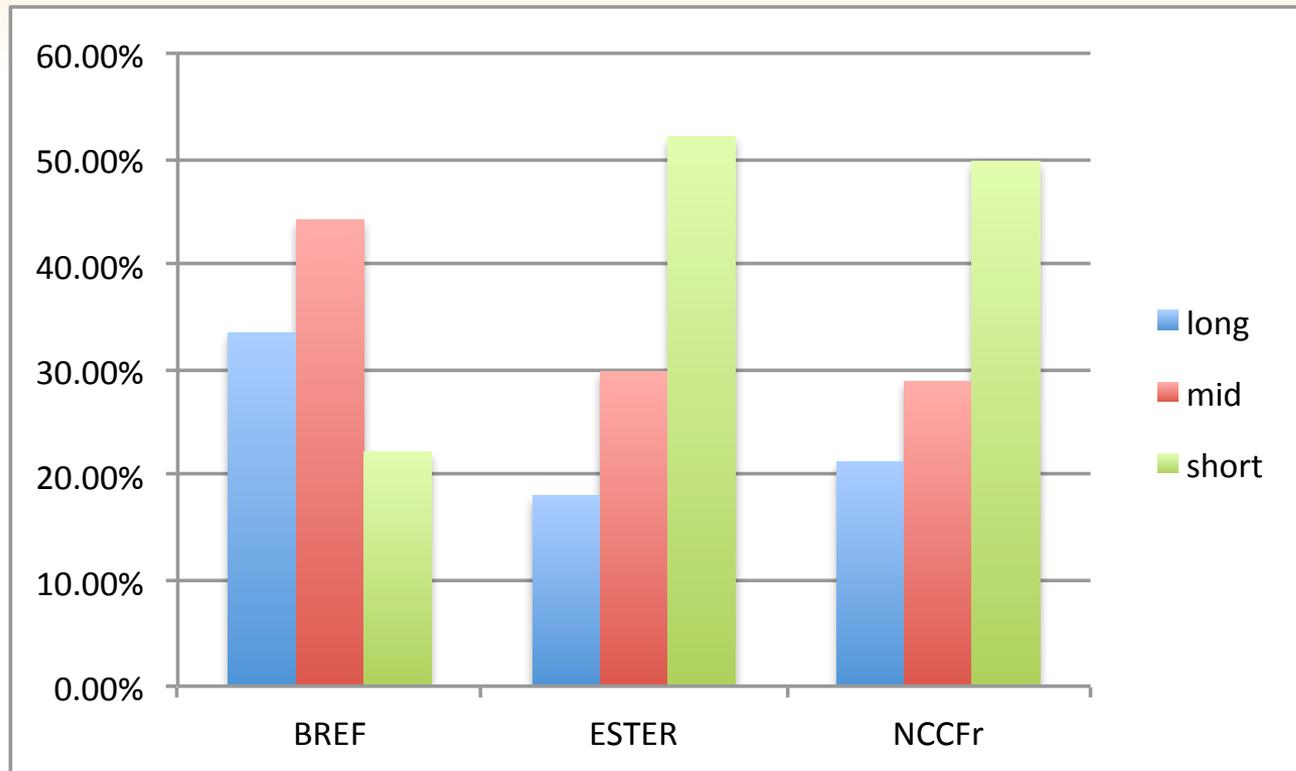
=> Mesures de distances entre voyelles (et donc métriques pour quantifier les variations de l'espace vocalique) plus conformes à la perception

# Distribution des voyelles par corpus



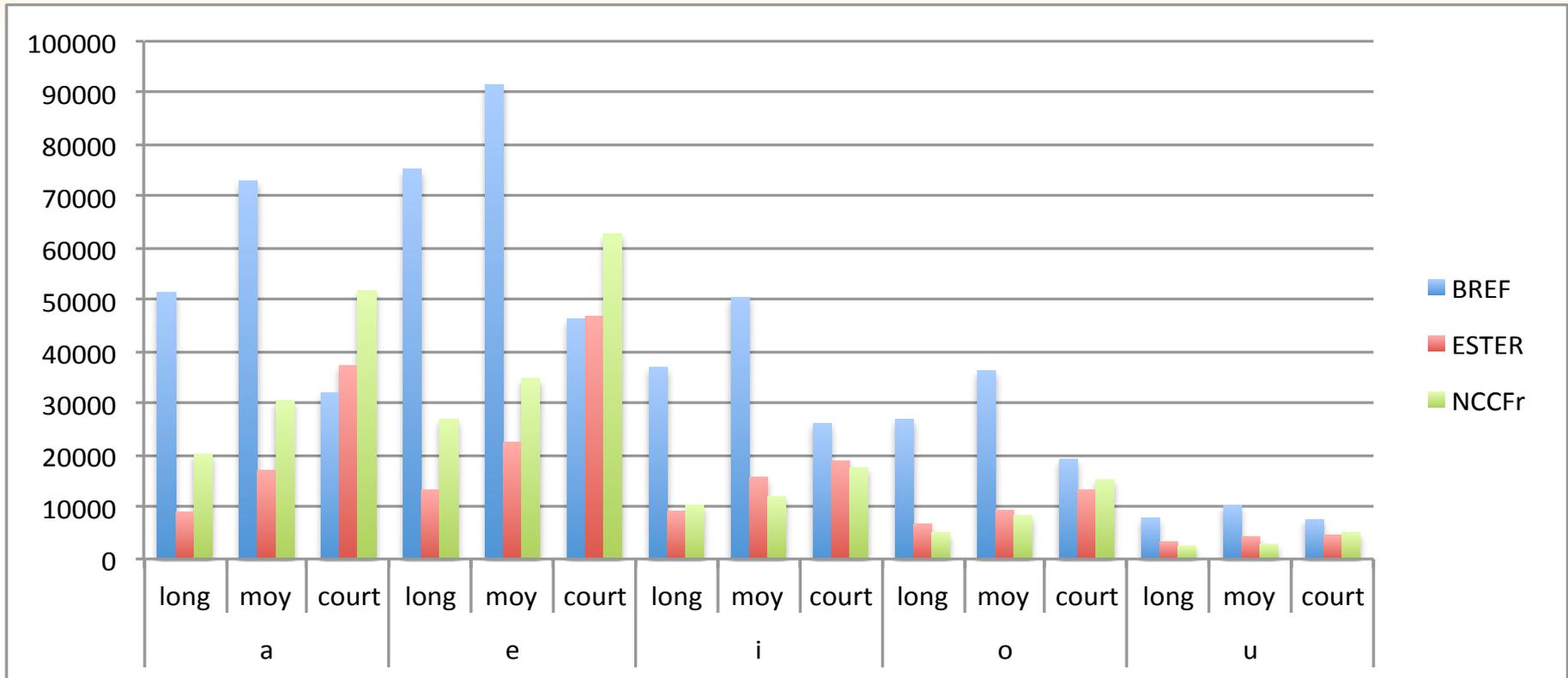
- Dans les trois corpus, sur-représentation de /a/ et /e/, sous-représentation de /u/

# Distribution des classes de durée par corpus



- Plus de voyelles moyennes et longues en parole lue

# Distribution des voyelles par corpus et classe de durée



- On retrouve l'influence du déséquilibre des classes de durée entre corpus

# Nombres d'exemplaires utilisés pour le calcul des métriques

- Élément de base pour le calcul des métriques :  
locuteur x classe\_durée  
(si mêmes locuteurs dans les 3 corpus,  
ce serait corpus x locuteur x classe\_durée)
- On cherche à obtenir des mesures représentatives du système vocalique dans son ensemble
  - ⇒ Nécessite une représentation fiable de la tendance générale et de la dispersion pour chaque voyelle
  - ⇒ Nombre minimum d'exemplaires par voyelle pour chaque locuteur x classe\_durée fixé à 15
  - ⇒ Élimination des locuteurs du corpus ESTER pour lesquels on ne dispose que de courts extraits

# Nombres d'exemplaires utilisés pour le calcul des métriques

| corpus | durée | #voy        | #a         | #e         | #i         | #o        | #u       |
|--------|-------|-------------|------------|------------|------------|-----------|----------|
| BREF   | long  | 1011 - 6910 | 270 - 1948 | 392 - 2519 | 169 - 1315 | 83 - 857  | 19 - 291 |
|        | moy   | 1658 - 5344 | 411 - 1569 | 573 - 1982 | 302 - 1086 | 242 - 791 | 50 - 258 |
|        | court | 325 - 4194  | 51 - 1167  | 100 - 1697 | 83 - 823   | 62 - 687  | 20 - 226 |
| ESTER  | long  | 171 - 2696  | 34 - 500   | 50 - 956   | 22 - 609   | 18 - 517  | 15 - 279 |
|        | moy   | 255 - 5264  | 52 - 1266  | 70 - 1780  | 38 - 1286  | 34 - 836  | 15 - 362 |
|        | court | 444 - 10903 | 111 - 3397 | 190 - 4182 | 55 - 1758  | 39 - 1212 | 15 - 354 |
| NCCFr  | long  | 491 - 3274  | 134 - 1219 | 166 - 1356 | 58 - 538   | 30 - 221  | 18 - 105 |
|        | moy   | 813 - 3703  | 242 - 1227 | 298 - 1638 | 106 - 670  | 62 - 356  | 20 - 122 |
|        | court | 1540 - 7017 | 516 - 2587 | 562 - 2750 | 181 - 761  | 134 - 733 | 42 - 246 |

- Systèmes vocaliques analysés comptent au total entre 171 et 10903 voyelles
- Malgré la taille des corpus, impossible de contrôler les contextes consonantiques sans éliminer la plupart des locuteurs

# Résultats : corrélations

- Corrélations faibles entre les trois dimensions de la réduction vocalique mesurées  
⇒ Confirme l'utilité d'une approche multidimensionnelle

|                | centralisation | dispersion | contraste |
|----------------|----------------|------------|-----------|
| centralisation |                | -0.195     | -0.304    |
| dispersion     |                |            | -0.195    |

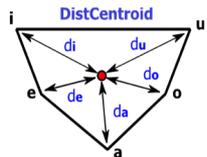
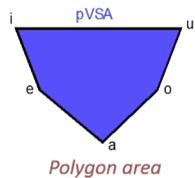
# Résultats :

## Centralisation / réduction

- Mêmes effets et magnitude comparable pour les 4 métriques

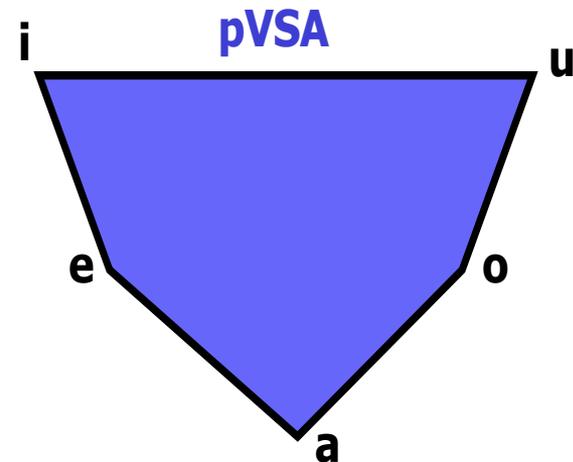
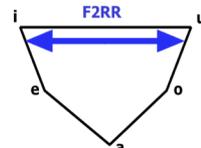
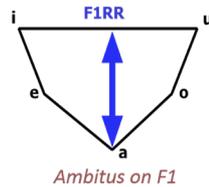
⇒ Focus sur l'aire du polygone

- Overall

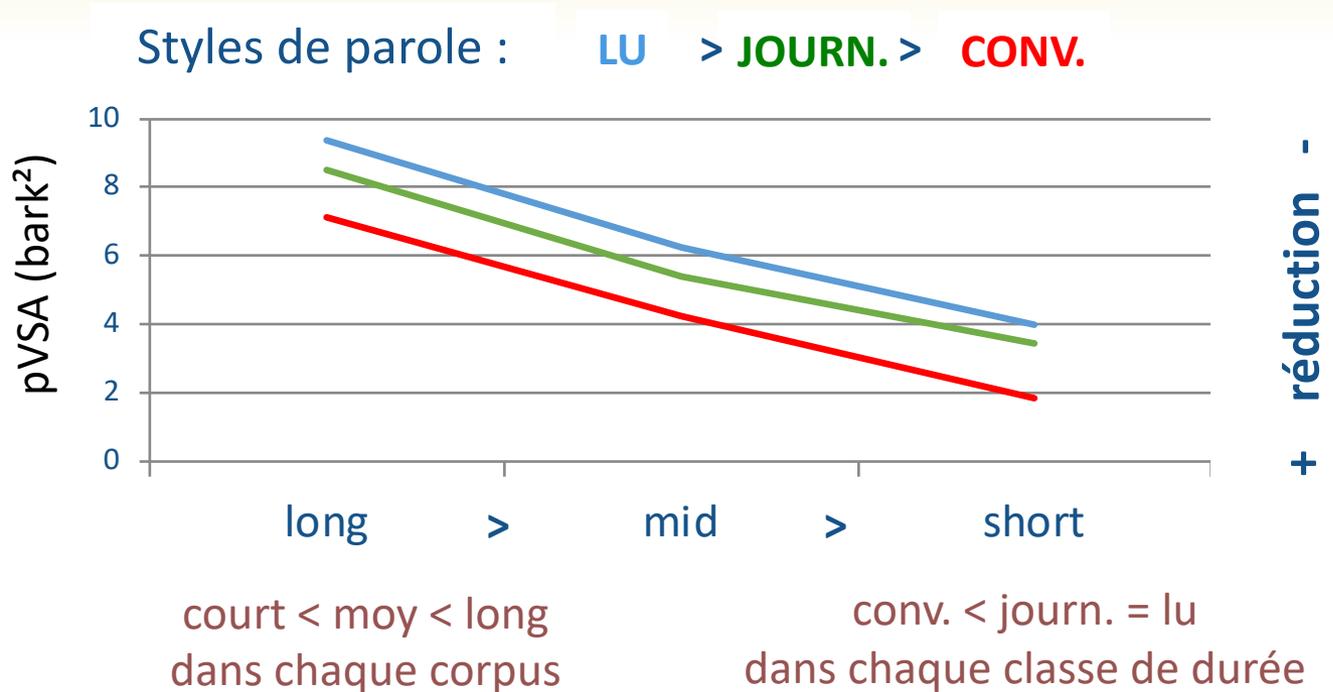
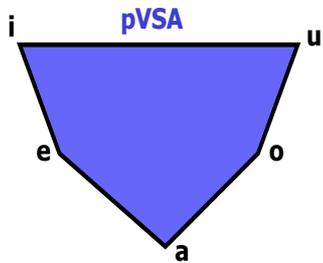


*Mean distance to vowel space centroid*

- Specific to an acoustic dimension (Sapir et al. 2010)

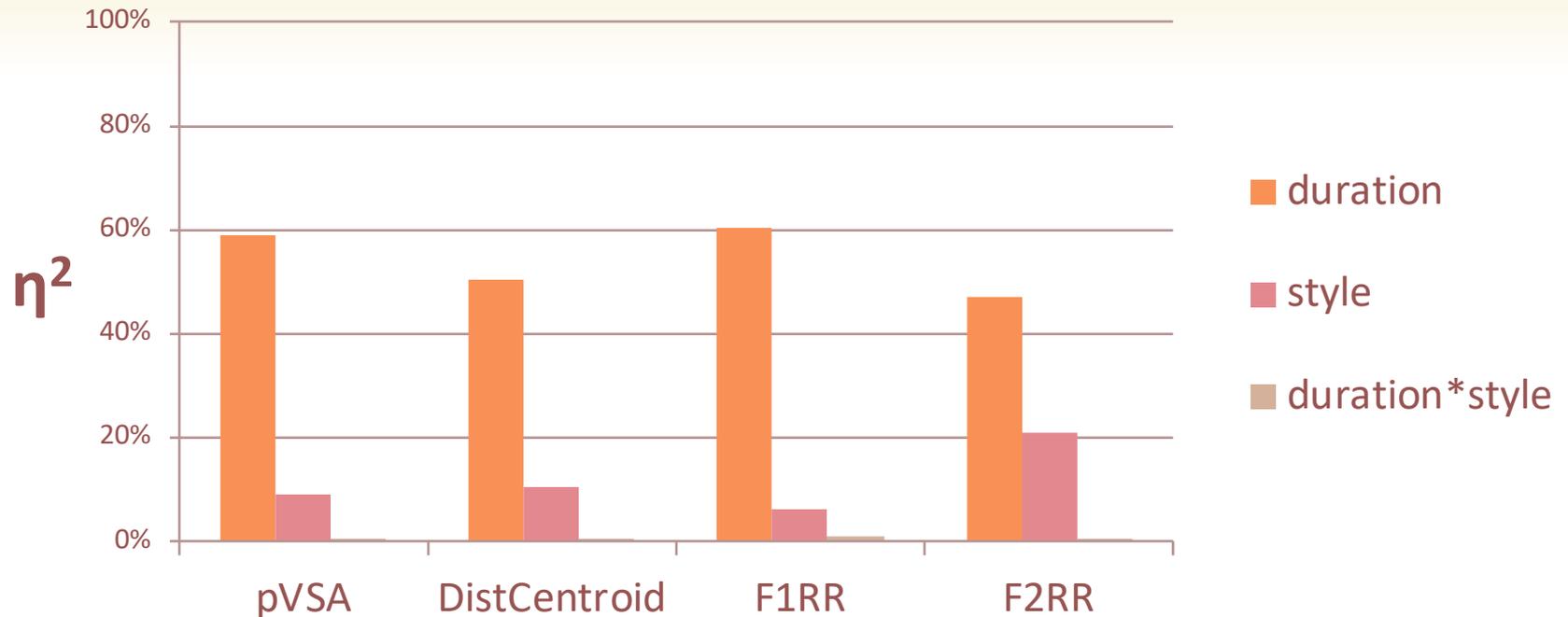


# Résultats : Centralisation / réduction



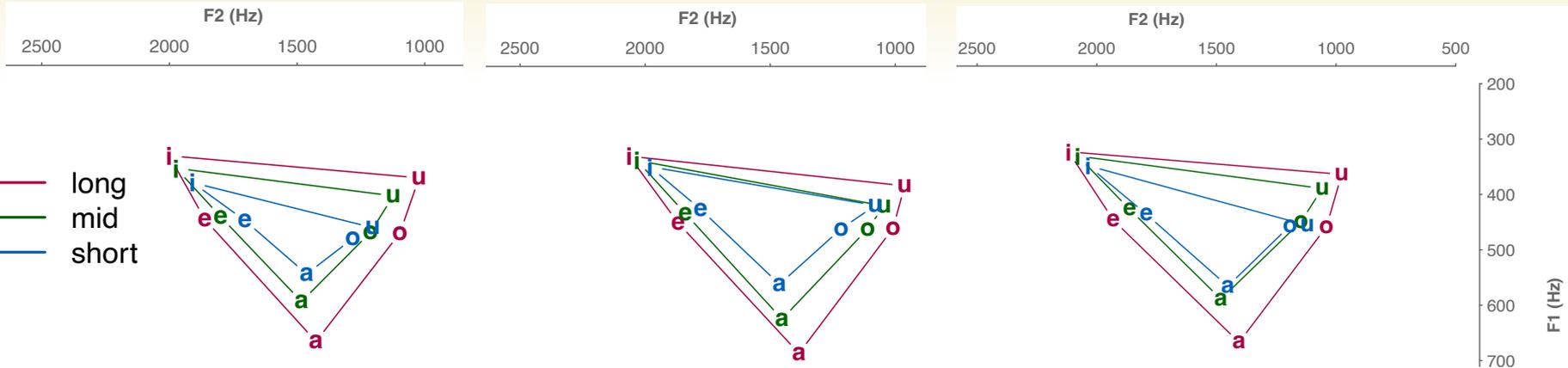
# Mesures de centralisation/réduction

## Comparaison des tailles d'effet



- DUREE : effet sur F1 range > F2 range
- STYLE : effet sur F2 range > F1 range

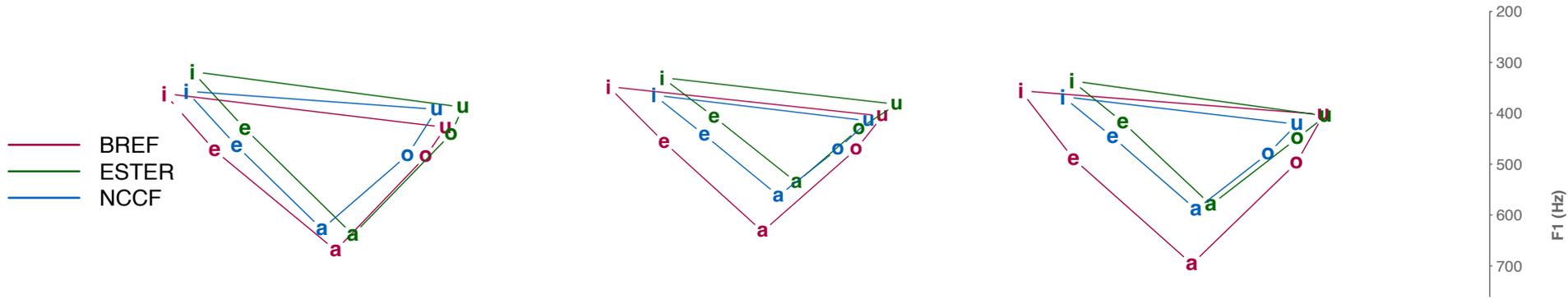
# Analyse qualitative des espaces vocaliques : différences entre corpus et classes de durée



Conversational speech: NCCFr

Journalistic speech: ESTER

Read speech: BREF



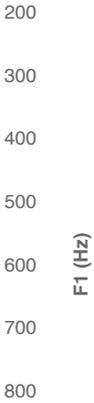
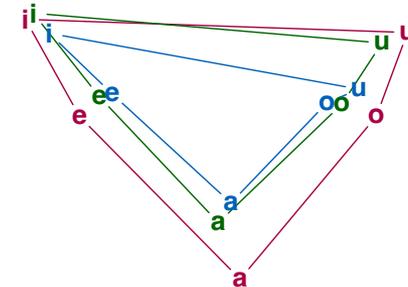
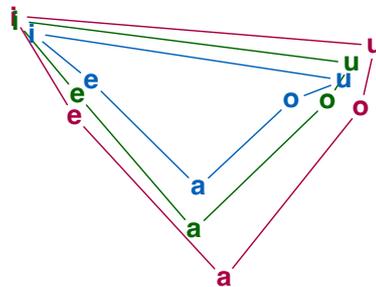
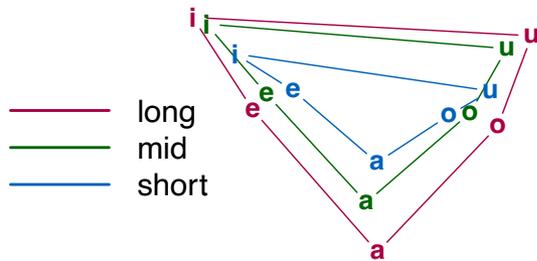
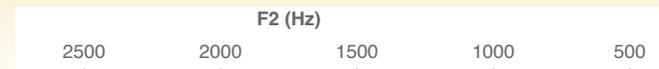
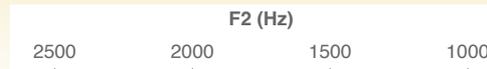
Short vowels (≤ 50ms)

Mid vowels (≤ 80ms)

Long vowels (> 80ms)

Male speakers

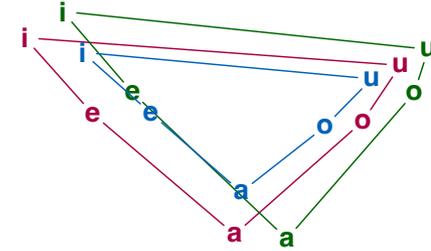
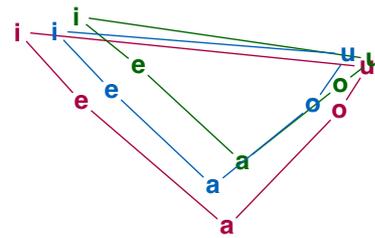
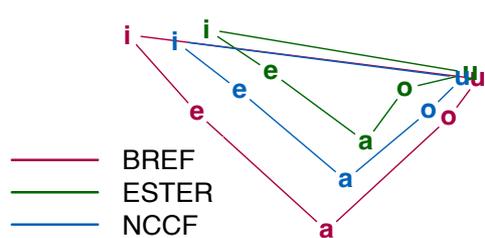
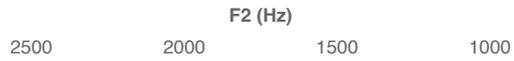
# Analyse qualitative des espaces vocaliques : différences entre corpus et classes de durée



Conversational speech: NCCFr

Journalistic speech: ESTER

Read speech: BREF



Short vowels ( $\leq 50$ ms)

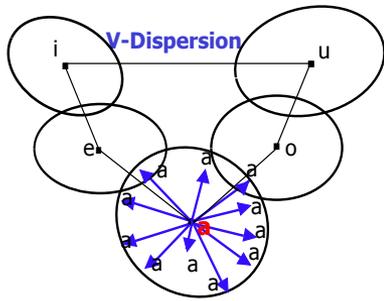
Mid vowels ( $\leq 80$ ms)

Long vowels ( $> 80$ ms)

Female speakers

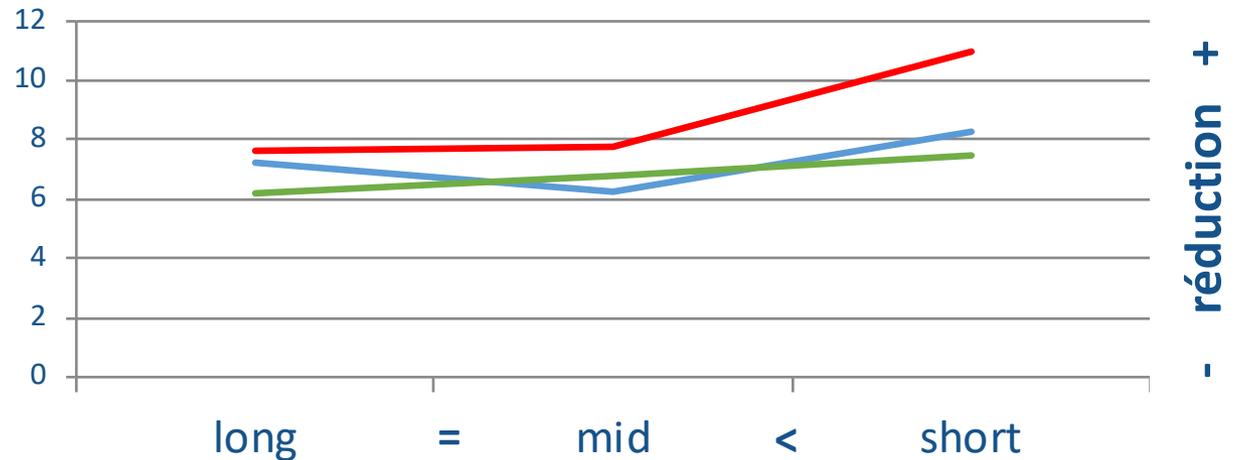
# Résultats :

## Dispersion intra-catégories



V-Dispersion (bark<sup>2</sup>)

Styles de parole : LU = JOURN. < CONV.



court < moy. < long  
dans chaque corpus

lu = journ. < conv.  
dans chaque classe durée

Pas de distinction entre :

- Style lu et journalistique
- Voyelles longues et moyennes

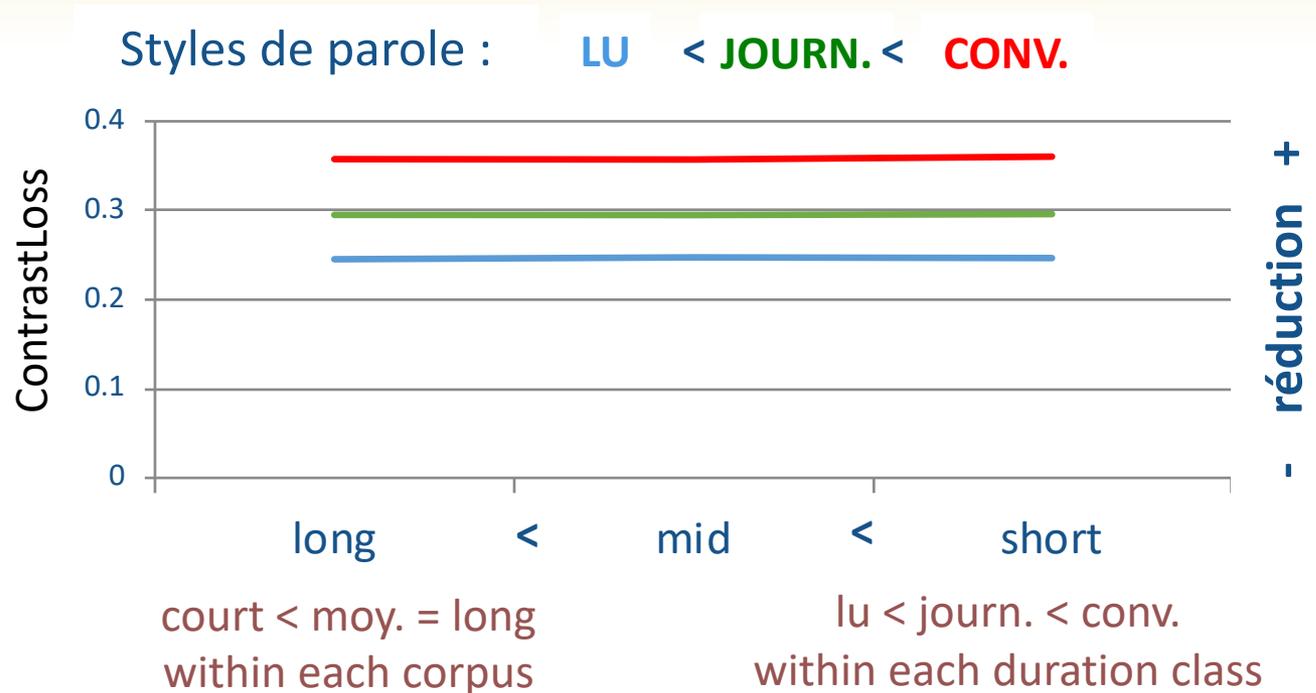
# Résultats :

## Recouvrement entre catégories

actual vowel class membership

|   | a | e | i | u | o |
|---|---|---|---|---|---|
| a |   |   |   |   |   |
| e |   |   |   |   |   |
| i |   |   |   |   |   |
| u |   |   |   |   |   |
| o |   |   |   |   |   |

predicted from F1/F2

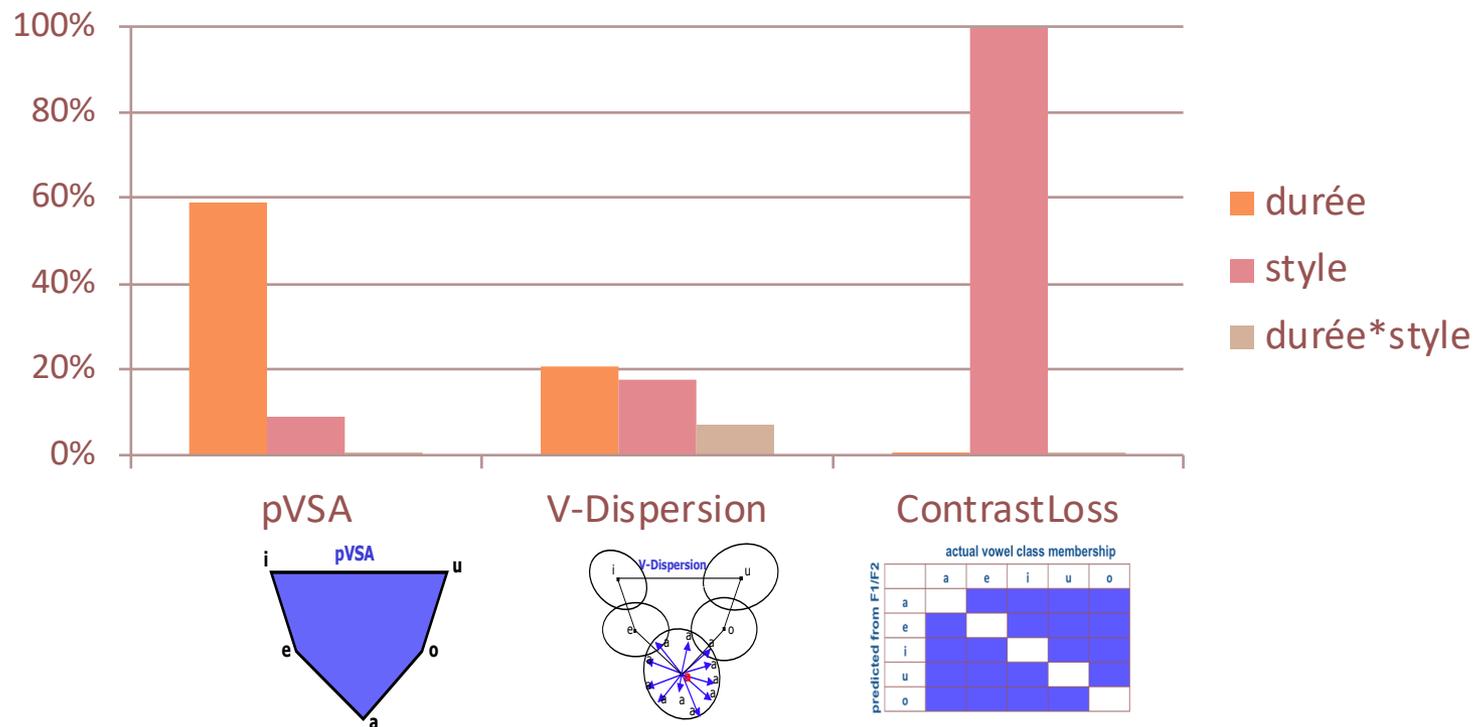


- Distinction très nette entre styles de parole
- Plus de variabilité pour /u/ (sans lien direct avec nb exemplaires,  $r=.19$ )
- Différences entre styles surtout pour /e/ et /o/, plus largement confondues lorsque le style est plus relâché

# Résultats :

## Comparaison des tailles d'effet

- Taille espace vocalique principalement liée à la durée
- Perte de contraste principalement liée au style
- Effets plus équilibrés pour la dispersion intra-classe

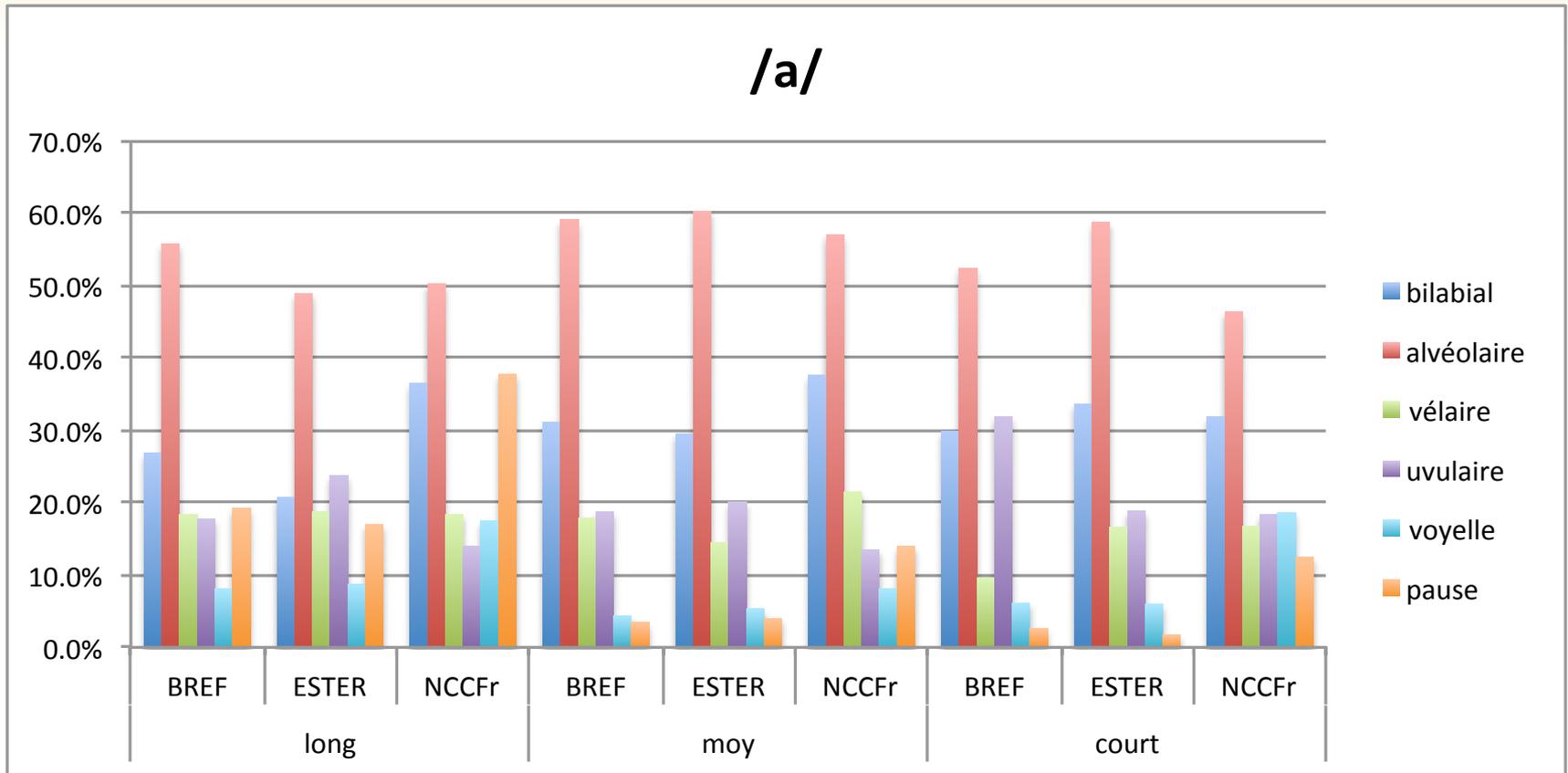


# Contextes consonantiques ? (1)

- Contextes consonantiques regroupés par lieu d'articulation, voyelles et pauses considérées séparément
- Distribution similaire des contextes gauche et droit, mais pas en miroir : il faut prendre en compte les contextes gauche et droit conjointement
- Illustration avec contexte droit pour la lisibilité des figures

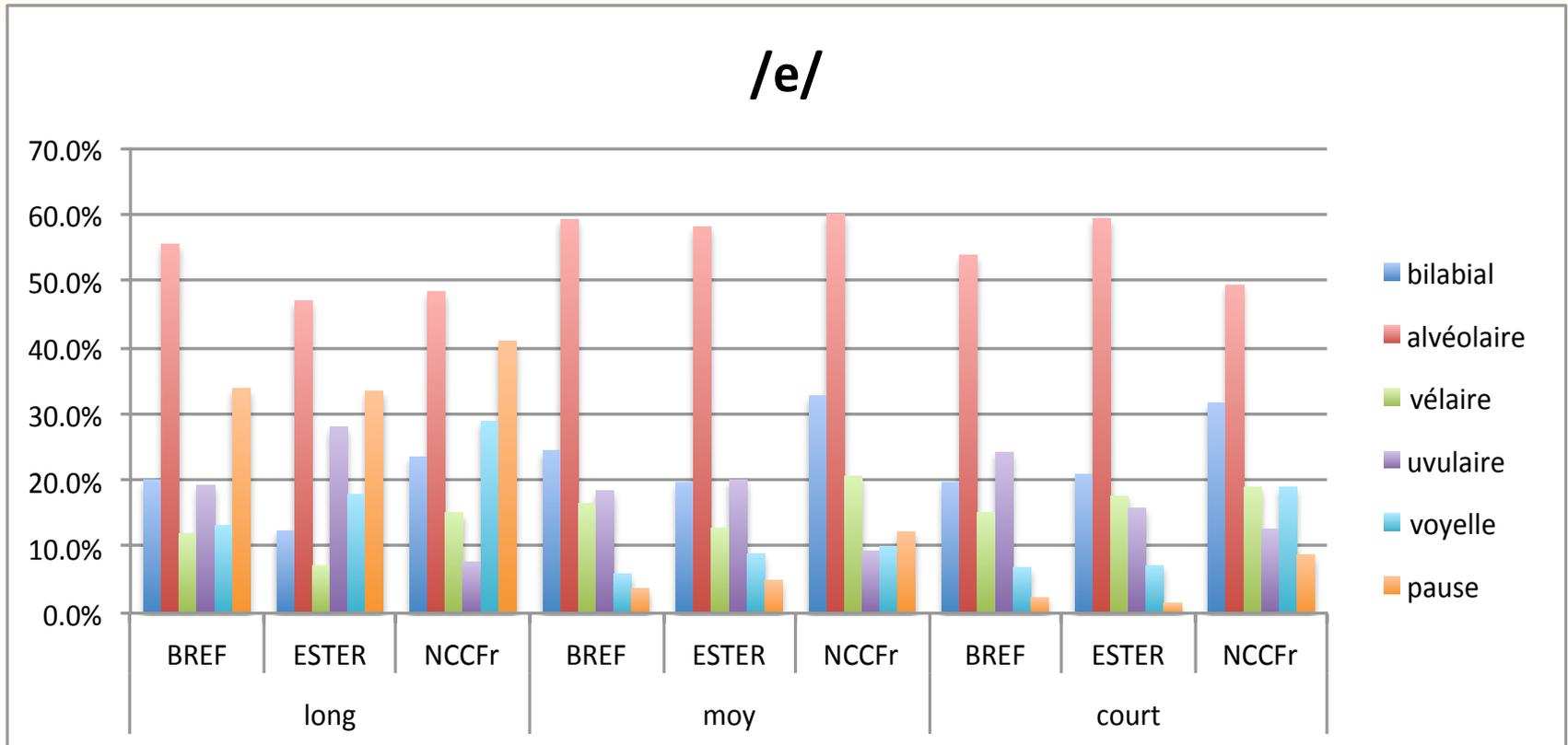
# Contextes consonantiques ? (2)

*/a/*



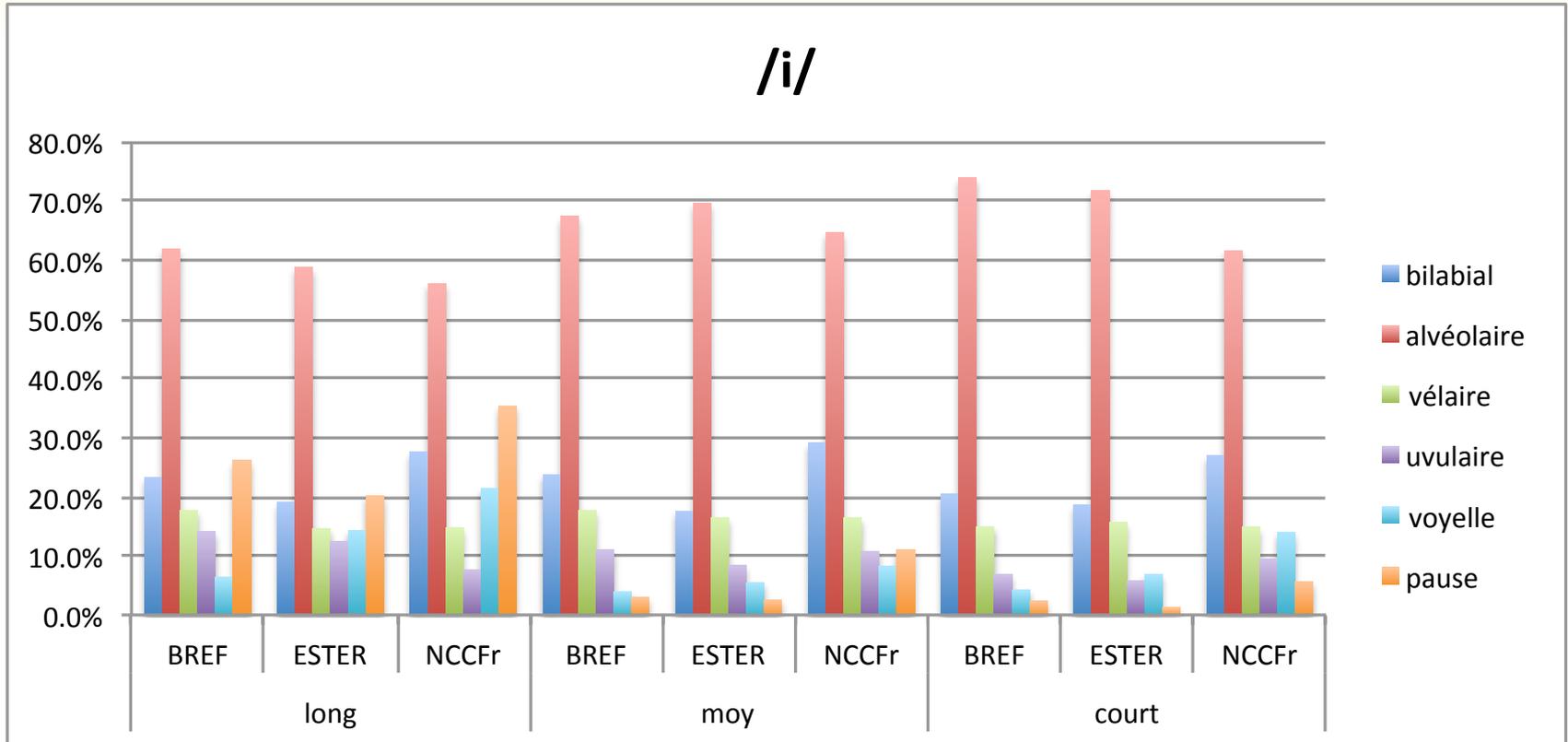
# Contextes consonantiques ? (3)

*/e/*



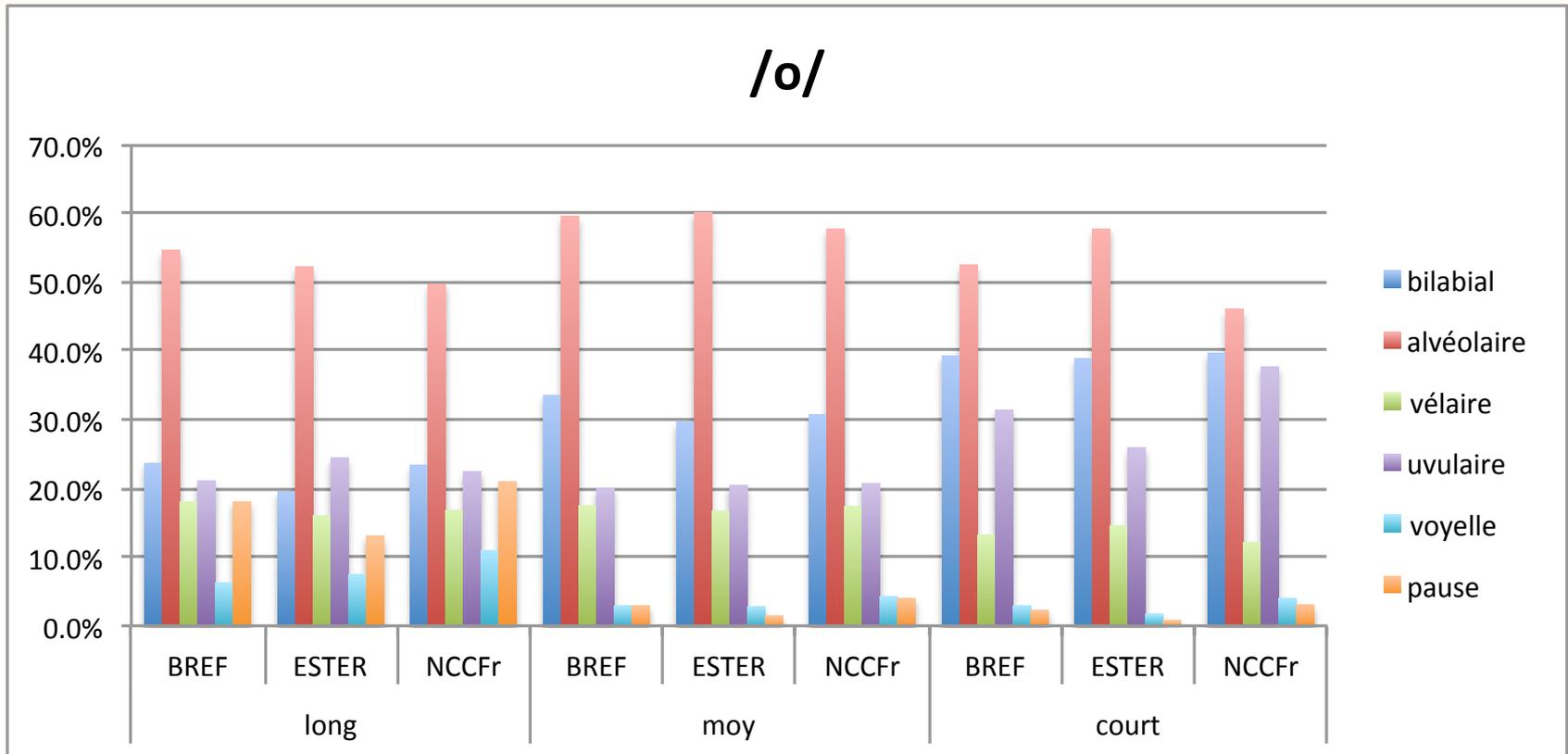
# Contextes consonantiques ? (4)

*/i/*

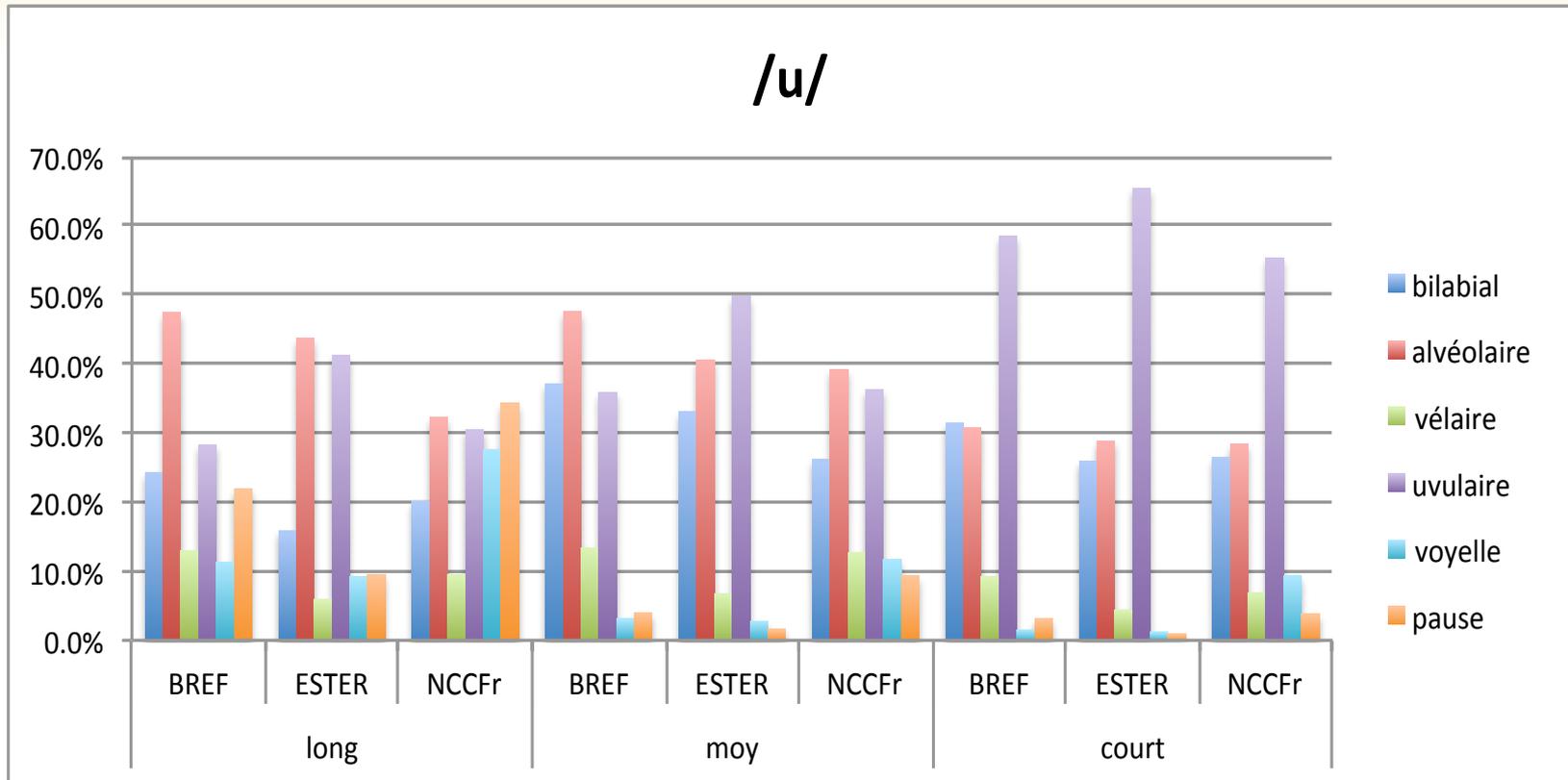


# Contextes consonantiques ? (5)

/o/



# Contextes consonantiques ? (6)



# Contextes consonantiques ? (7)

- Différences de distribution observées entre classes de durée et dans une moindre mesure entre corpus  
⇒ Peu susceptibles d'expliquer seules les effets mesurés
- Voyelles généralement plus longues quand suivies d'une pause (proportion > dans la classe *long*)
- Distributions comparables pour /a, e, i, o/ : contextes alvéolaires /t, d, n/ prépondérants, suivis par les contextes bilabiaux /p, b, m/
- Différent pour /u/ avec surreprésentation des contextes uvulaires (/ʁ/), particulièrement pour voyelles courtes. En partie dû aux nombreuses occurrences du mot *pour*.

# Conclusion (1)

1. Contribution des facteurs durée et style à la réduction vocalique : réduction + forte pour voyelles courtes en parole conversationnelle
  - Espace vocalique de taille réduite
  - Cibles acoustiques plus instables : dispersion intra-catégorie plus importante
  - Réduction des contrastes entre catégories : perte de précision dans la catégorisation à partir de F1 et F2

# Conclusion (2)

2. La réduction ne peut se réduire à la taille de l'espace vocalique : confirme la nécessité d'une approche multidimensionnelle  
(Fougeron & Audibert, 2011 ; Audibert & Fougeron, 2012)
3. Différences de style
  - parole conversationnelle ≠ autres styles
  - lecture ≈ journalistique sauf pour perte contraste

# Distinction /e/ vs. /ɛ/ en position finale

Gendrot, C., Audibert, N. (2019). *La distinction /e/ vs. /ɛ/ est-elle toujours produite en français standard ? Etude sur des corpus de parole journalistique et de parole spontanée*. Langue Française n°203(3), 53-65.

# Distinction /e/ vs. /ɛ/ en position finale

- Processus en cours de fusion entre les voyelles moyennes /e/ et /ɛ/ en français ?
  - Variable entre styles de parole ?
- ⇒ Pour 157 locuteurs des corpus ESTER et NCCFr, mesure de la distance euclidienne dans l'espace F1, F2 (en Bark) entre :
- /e/ et /ɛ/
  - /ɛ/ et /a/ à titre de comparaison

# Méthodologie

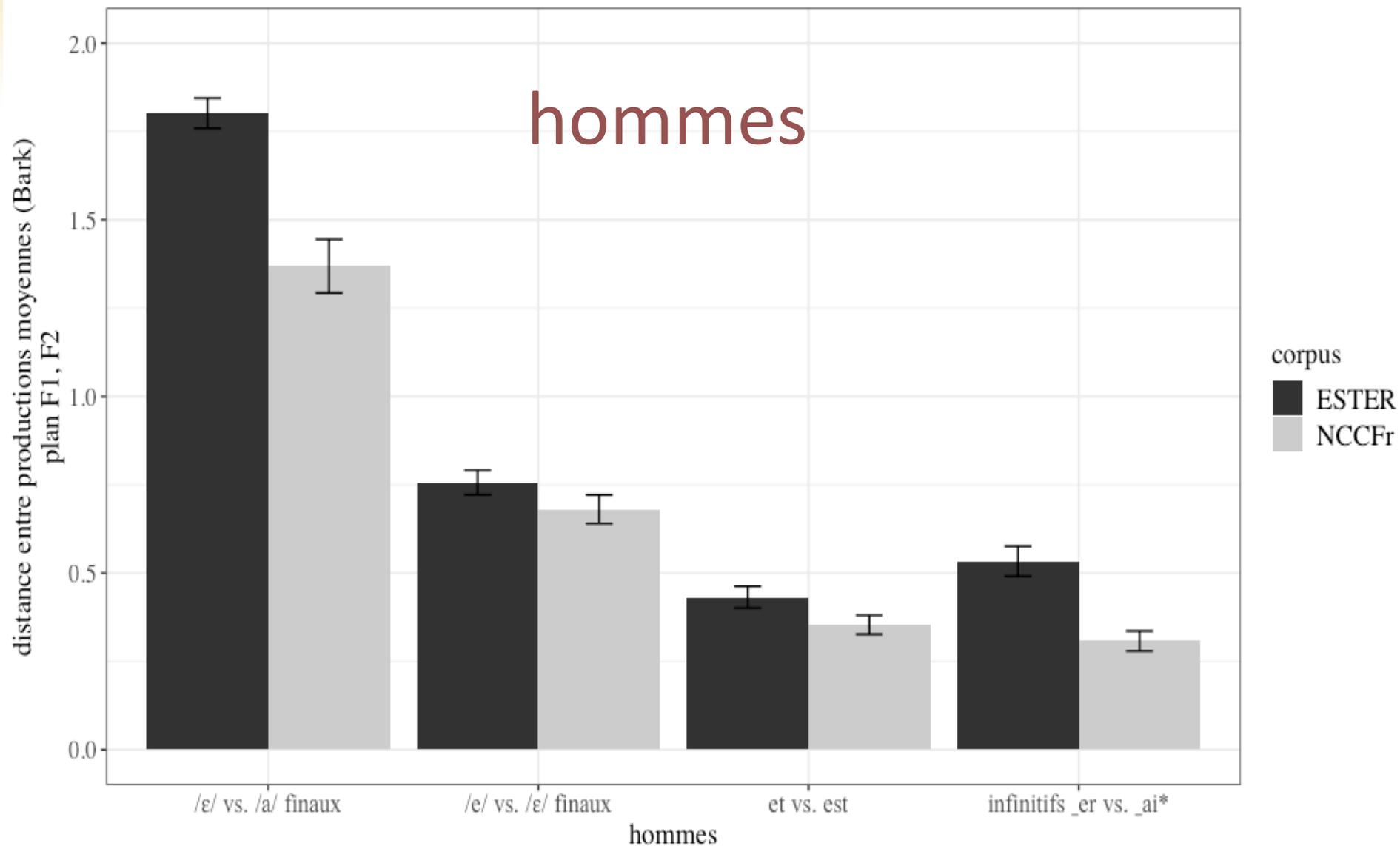
- Méthode similaire à celle présentée précédemment pour l'extraction des formants (détection auto. + filtrage via crible)
- Hormis l'analyse spécifique sur et vs. est, élimination de mots grammaticaux très fréquents (21% des occurrences des voyelles ciblées)
- Seuil pour l'inclusion des locuteurs fixé à 15 occurrences de chacune des voyelles

# Conditions prises en compte

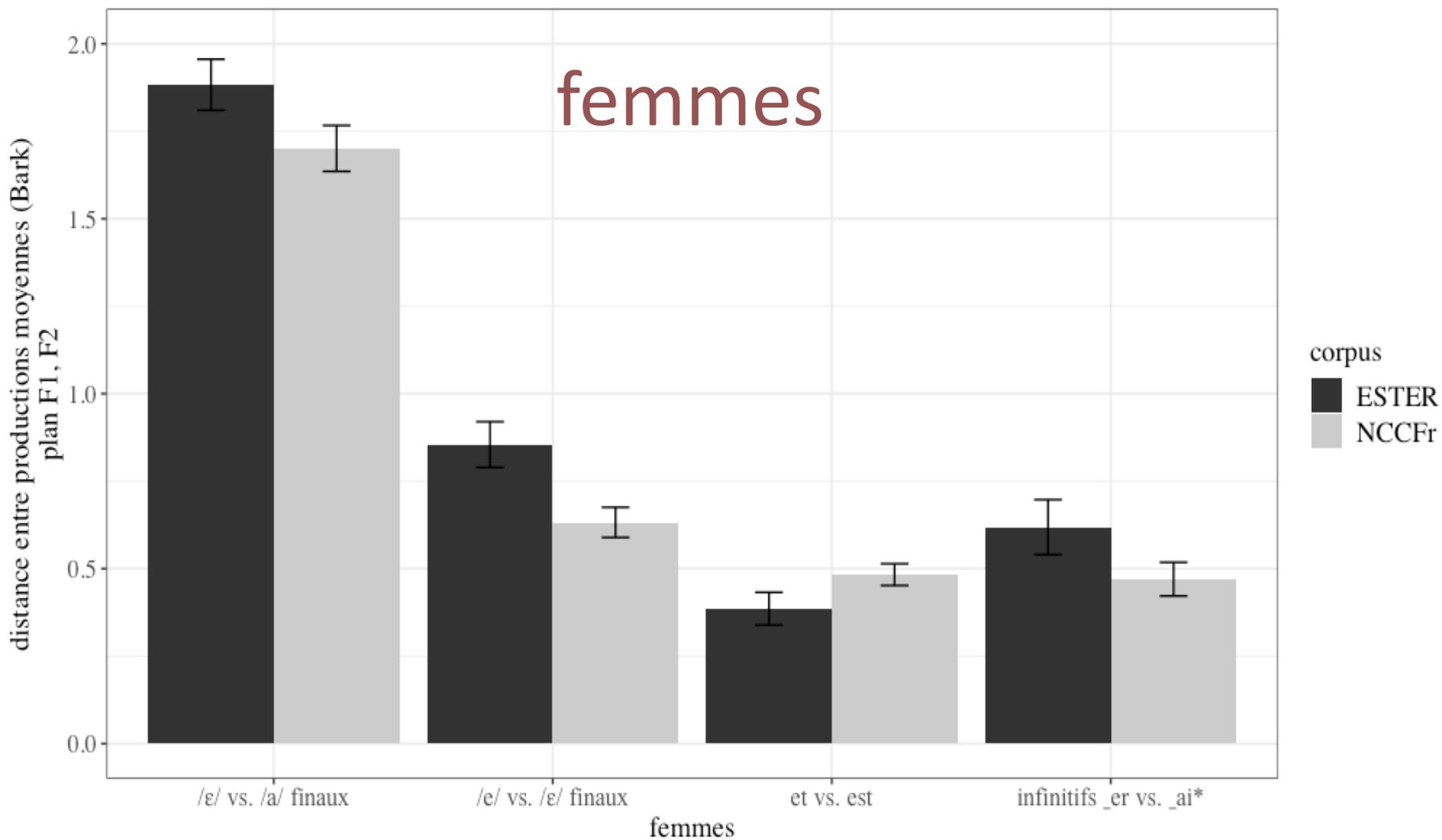
- Distance entre /ε/ et /a/ en position finale de mot (120k voyelles, dont 72 % correspondent à des paires minimales)
- Distance entre /e/ et /ε/ en position finale de mot (76k voyelles)
- Distance entre /e/ et /ε/ pour l'opposition entre *et* et *est* seulement (29k voyelles)
- Distance entre /e/ et /ε/ pour l'opposition supposée entre l'infinitif en «-er» et l'imparfait ou conditionnel «-ais», «-ait» ou «-aient» (23k voyelles)

# Résultats

hommes



# Résultats



# Résultats

- Dans les deux styles de parole, distance /e-ε/ < /ε-a/
- ⇒ Suggère une tendance à la neutralisation de l'opposition /e-ε/
- Réduction de la distance /e-ε/ plus importante pour *et* vs. *est*, et dans une moindre pour la distinction entre infinitif et formes en ai\*

# Discussion

- Nombreuses questions restantes
  - Idem pour opposition / $\emptyset$ - $\text{œ}$ / et / $\text{o}$ - $\text{ɔ}$ / ?
  - La distance / $\text{i}$ - $\text{e}$ / (non prise en compte car nb occurrences insuffisant) plutôt que / $\text{ɛ}$ - $\text{a}$ / comme référence aboutirait-elle au même résultat ?
  - Influence du contexte segmental ?

# Effet de l'âge sur l'espace vocalique

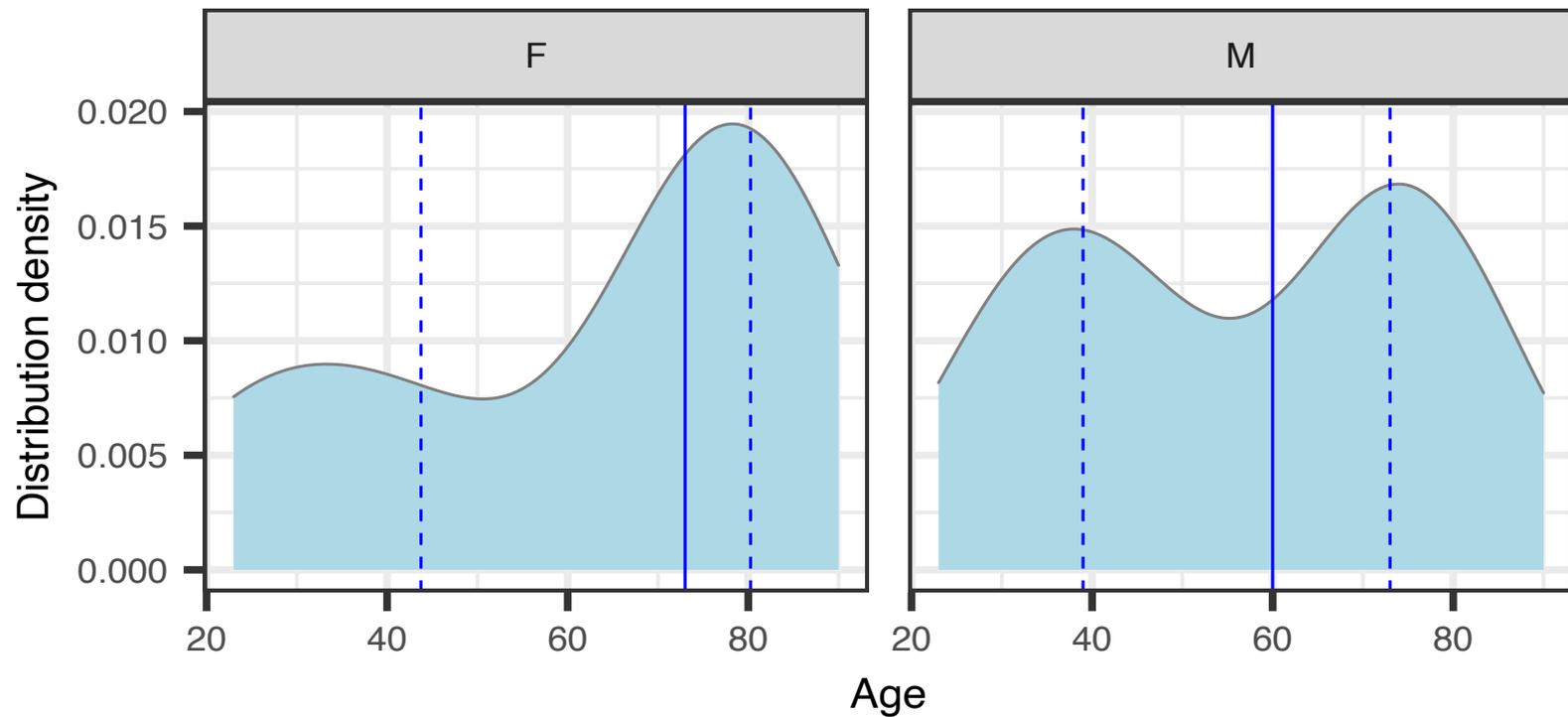
Hermes, A., Audibert, N., & Bourbon, A. (2023, August). AGE-RELATED VOWEL VARIATION IN FRENCH. In *20th International Congress of Phonetic Sciences* (pp. 2045-2049)

# Effet de l'âge sur la parole

- Résultat le plus robuste : baisse du débit
- Baisse F0 pour femmes, moins clair pour hommes
- VOT plus court pour les occlusives sourdes
- Résultat divergents sur les voyelles, hormis une baisse de F1 pour les femmes
- Pas de tendance claire en ce qui concerne la dispersion/centralisation, peu de données sur la réalisation des contrastes

# Méthodologie

- 37 locuteurs francophones âgés de 23 à 90 ans

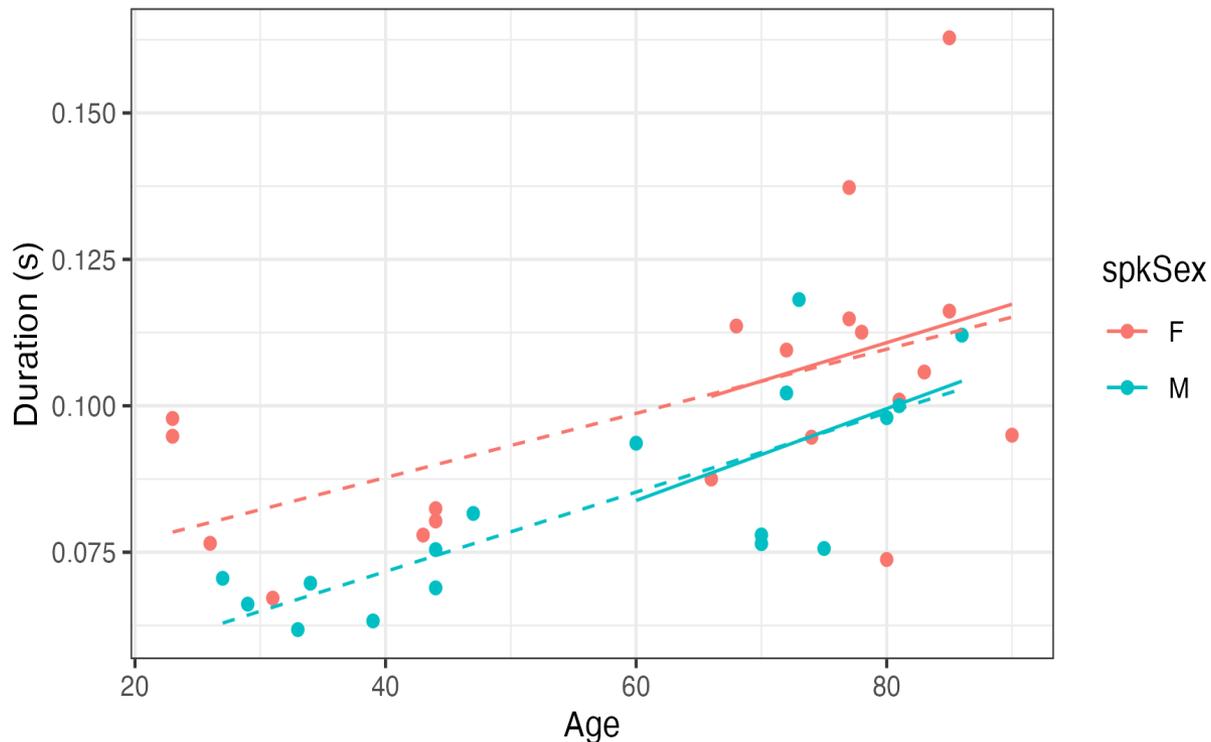


# Méthodologie

- Matériel contrôlé : mêmes phrases lues par tous les locuteurs
- Inclusion de 8 catégories vocaliques dont 2 voyelles nasales :  
*/i/, /e/, /ɛ/, /a/, /o/, /ɤ/, /õ/, /ã/*
- 15 375 voyelles au total
- Alignement forcé (WebMAUS) + correction manuelle

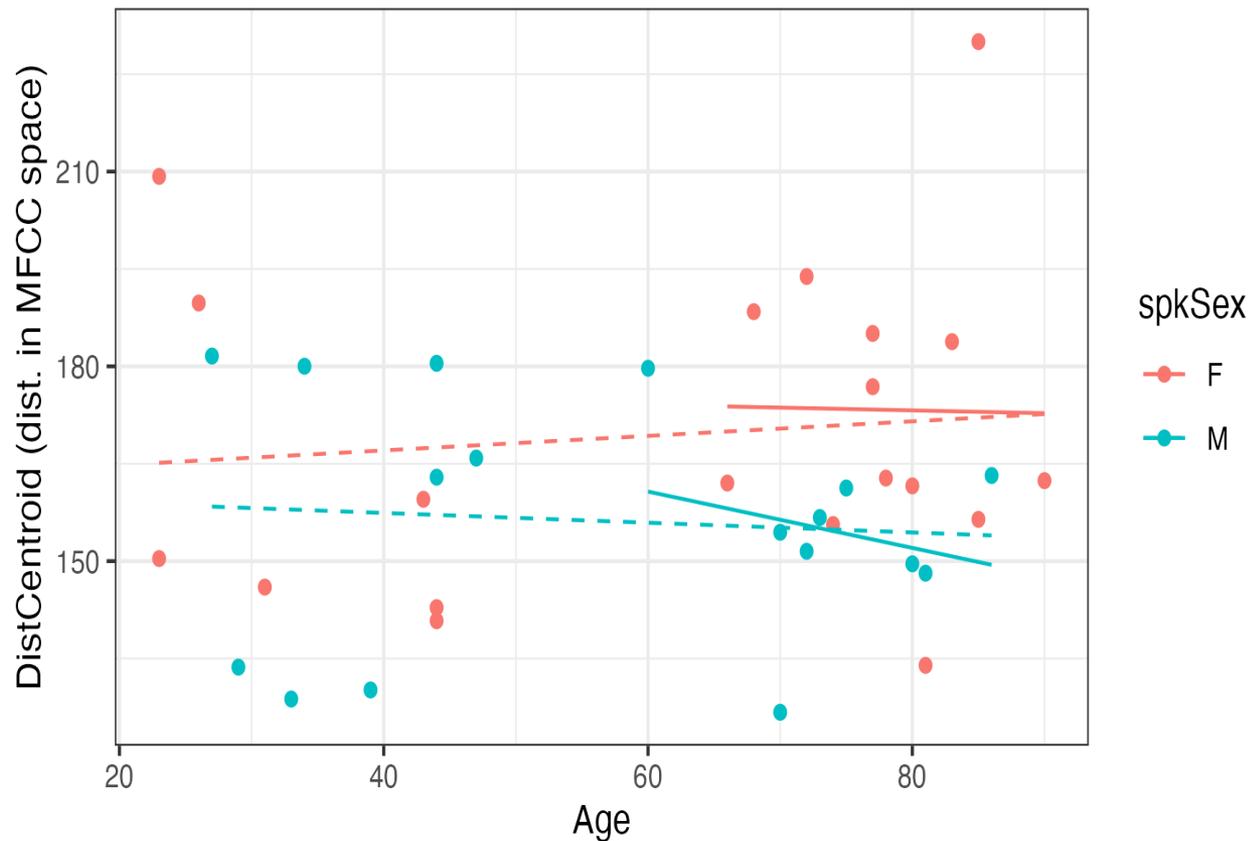
# Principaux résultats

- Effet de l'âge sur le débit de parole (ici durée segmentale) confirmé



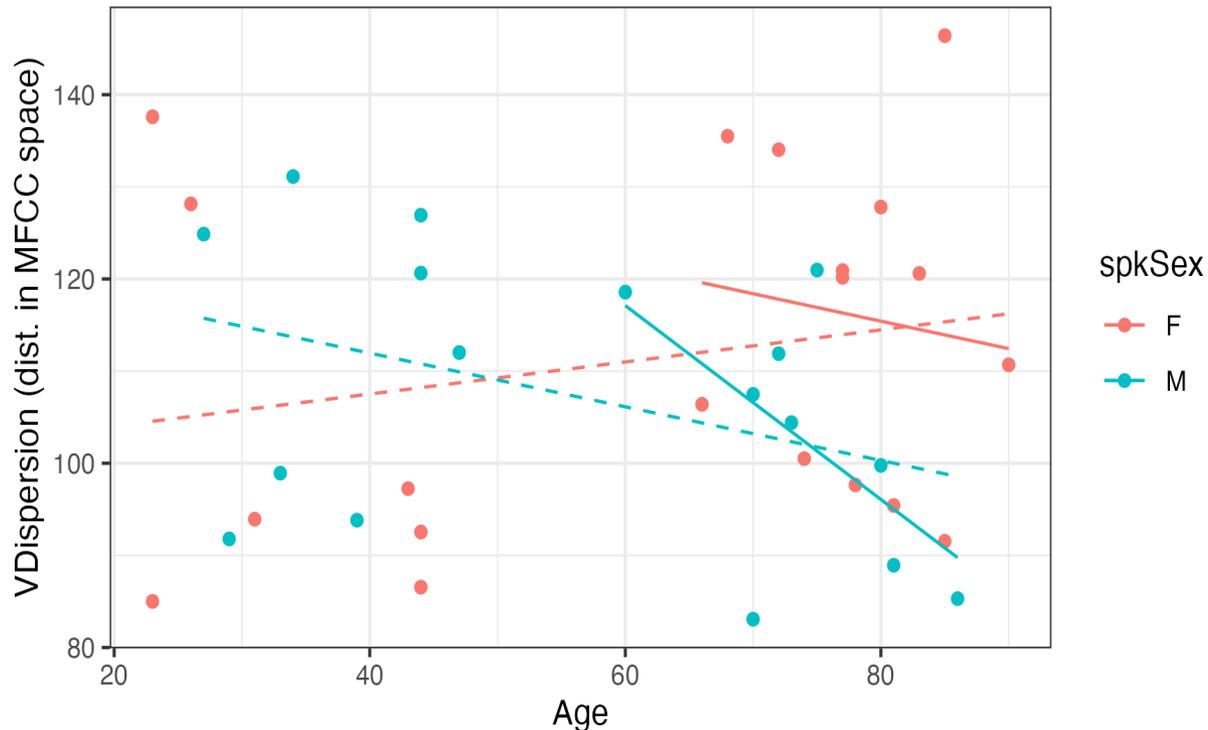
# Principaux résultats

- Pas d'effet clair sur la dispersion/centralisation



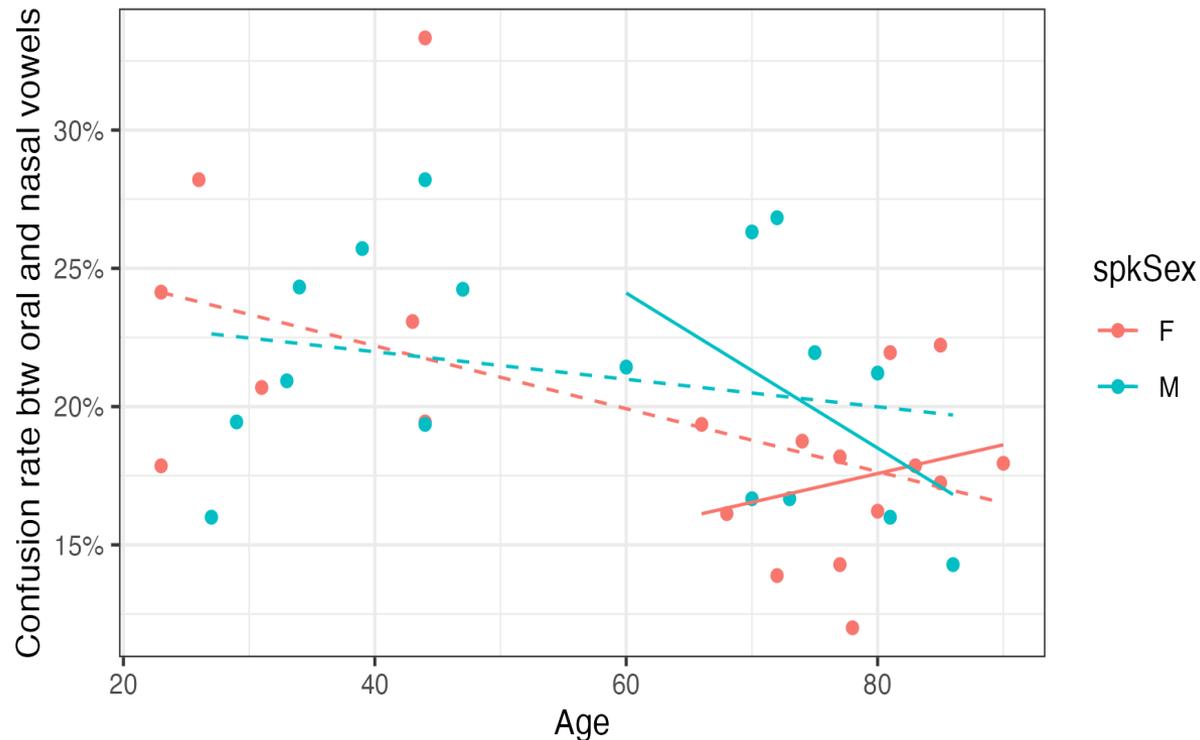
# Principaux résultats

- Tendances divergentes entre hommes et femmes pour la dispersion intra-catégorie : diminue avec l'âge surtout chez hommes âgés



# Principaux résultats

- Distinction entre orales et nasales augmente légèrement avec l'âge, débute plus jeune pour les femmes d'où divergence observée pour loc. âgés



# Discussion

- Comme dans toute étude « cross-sectional » (vs. longitudinale) sur les liens entre âge et parole, difficile de trancher entre effet de l'âge et différence générationnelle
- Avantages/inconvénients de l'utilisation des MFCC : permet prise en compte plus large du système vocalique et évite erreurs détection auto, mais perte interprétabilité directe

# Harmonie vocalique

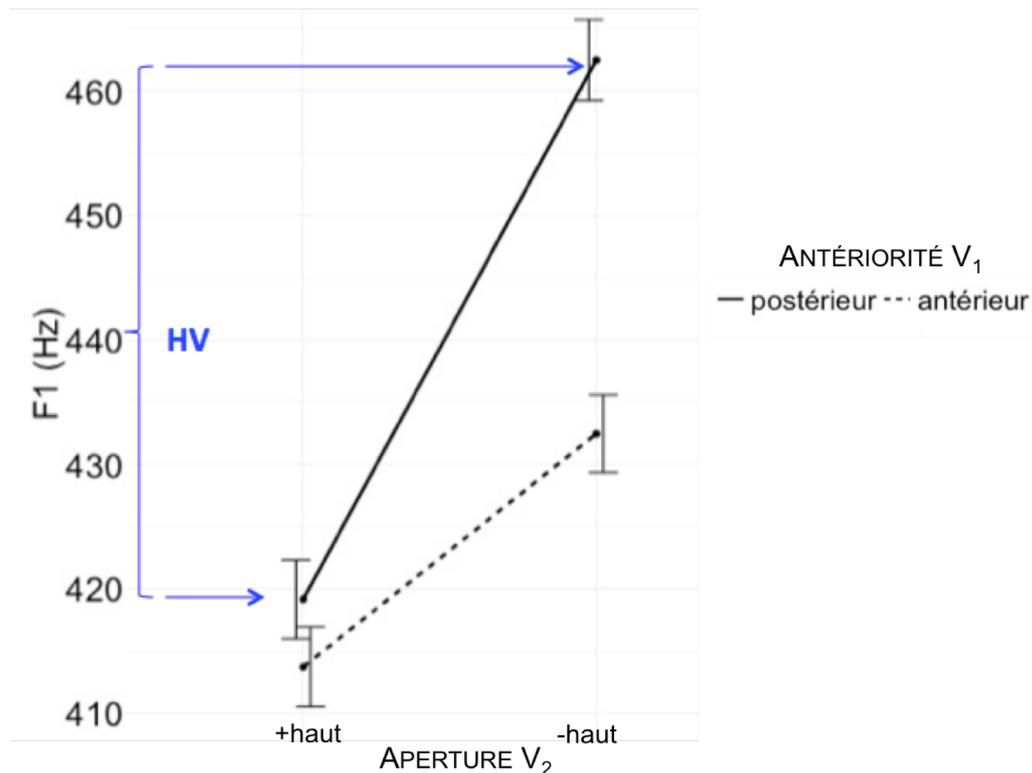
Turco, G., Fougeron, C., & Audibert, N. (2016, September). The effects of prosody on French V-to-V coarticulation: A corpus-based study. In *Interspeech* (pp. 998-1001).

# Harmonie vocalique en français

- Etude de l'harmonie vocalique en français : tendance à la modification du degré d'aperture des voyelles moyennes (mi-ouvertes et mi-fermées) en fonction de celui de la voyelle suivante
- Exemple : *aimait* [ɛmɛ] / *aimer* [eme]
- Objectif : mieux comprendre où apparaît l'harmonie vocalique, et dans quelle mesure les facteurs décrits dans la littérature la conditionnent

# Mesurer l'harmonie vocalique

- Mesure retenue de l'harmonie vocalique : différence de F1 sur V1 entre une V2 ouverte et une V2 fermée
- Evaluation sur 33k mots (ESTER 19k – NCCFr 14k)



# Principaux résultats

- Plus d'harmonie vocalique quand :
  - La 1<sup>e</sup> voyelle V1 est postérieure
  - Une graphie telle que « è » ou « au/eau » favorise une prononciation mi-fermée
  - V1 et V2 séparées par une consonne labiale (/p, b, m, f, v/) dont l'articulation n'implique pas la langue, plutôt que linguale
  - V1 et V2 ne sont pas séparées par un schwa sous-jacent (« e muet » non prononcé mais supposé appartenir à la représentation du mot)

# CONSONNES FRICATIVES

# Consonnes fricatives

- Produites par un flux d'air turbulent
- Corrélats acoustiques du lieu d'articulation des consonnes fricatives : moments spectraux, tout particulièrement le 1<sup>er</sup> (centre de gravité spectral COG), et dans une moindre mesure le 2<sup>e</sup> (dispersion spectrale)
- Dans le cas des sonores, mesures influencées par le voisement (énergie en BF qui tire les valeurs de COG vers le bas et augmente la dispersion, mais pas toujours dans les mêmes proportions)

# Réalisation comme approximantes des fricatives sonores

Dong S., & Audibert, N., en préparation

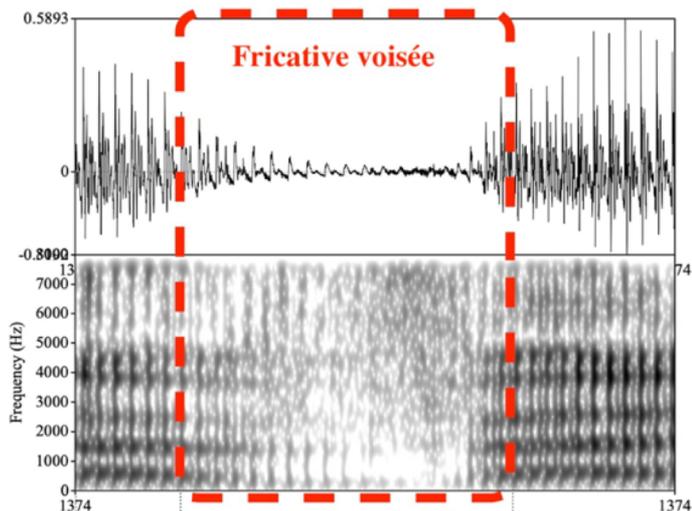
Résultats issus du mémoire  
de M1 de Suyuan DONG

# Réalisation comme approximantes des fricatives sonores

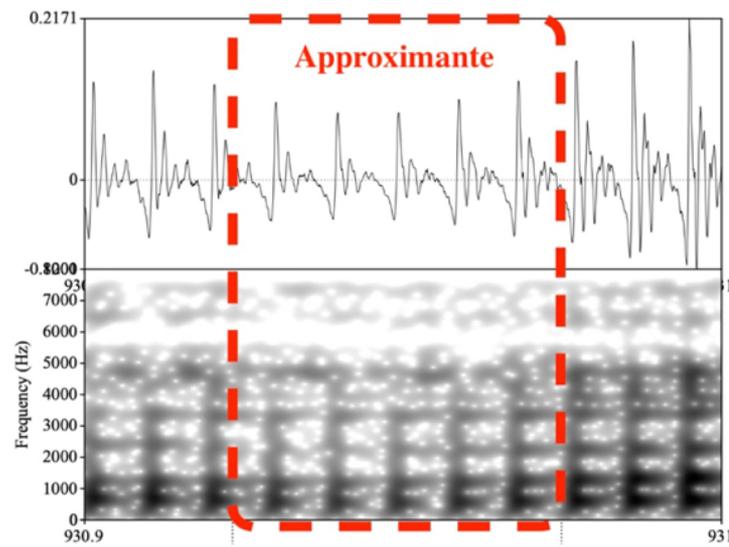
- Focus sur /v/ en position intervocalique (dans la parole spontanée, corpus NCCFr)
- Susceptible d'être réalisée comme [ʋ] (approximante labiodentale)
- Beaucoup moins étudié en français que le dévoisement (ex. Jatteau et al.)

# Réalisation comme approximantes des fricatives sonores

- Caractéristiques acoustiques attendues dans le cas d'une approximante :
  - Structure formantique similaire à voyelle, mais moins intense et moins stable
  - Energie acoustique < voyelle mais > fricative
  - Disparition bruit de friction



nombreuses  
possibilités de  
réalisations  
intermédiaires



# Méthodologie

- Présélection à partir de la segmentation disponible de toutes les occurrences de /v/ intervocaliques dans le corpus NCCFr
- Catégorisation (fricative ou approximante + critères secondaires) à partir de l'observation des spectrogrammes et d'une grille de critères
- Première étape : sur 10 hommes, 1728 occurrences exploitables

# Méthodologie

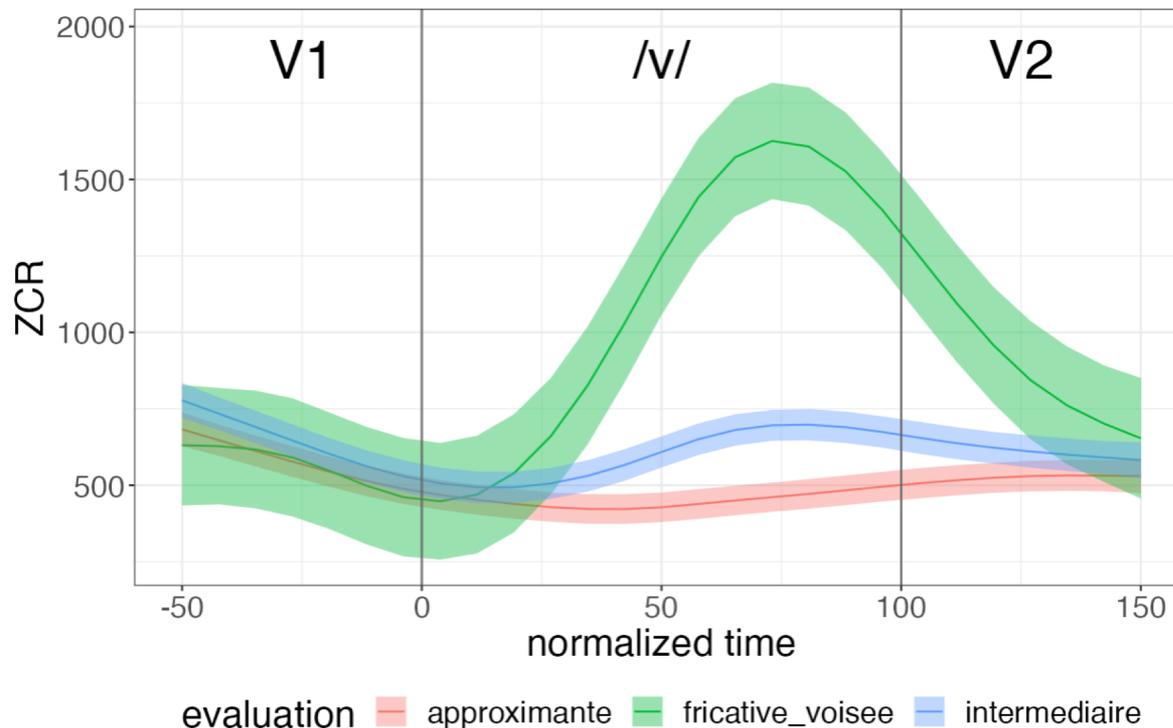
- Mesures acoustiques sur /v/ et les voyelles précédente et suivante :
  - Intensité dans différentes bande de fréquence
  - HNR
  - ZCR
  - COG
  - Mesures formantiques
- Modélisation GAM pour rendre compte de la dynamique des réalisations, prise en compte de la durée et du contexte segmental comme facteurs aléatoires

# Premiers résultats

- Sans surprise, réalisations approximantes plus fréquentes dans des contextes susceptibles de favoriser la réduction : segments plus brefs, positions prosodiques faibles
- Mais lien moins systématique qu'attendu, et variations entre locuteurs

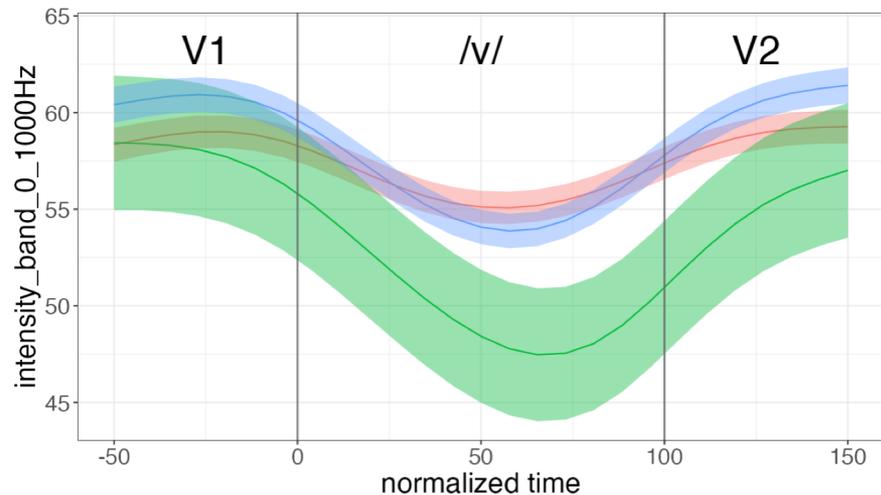
# Premiers résultats

- Mesures acoustiques qui discriminent bien les deux types de réalisation :
  - Mesures de périodicité (HNR et ZCR)

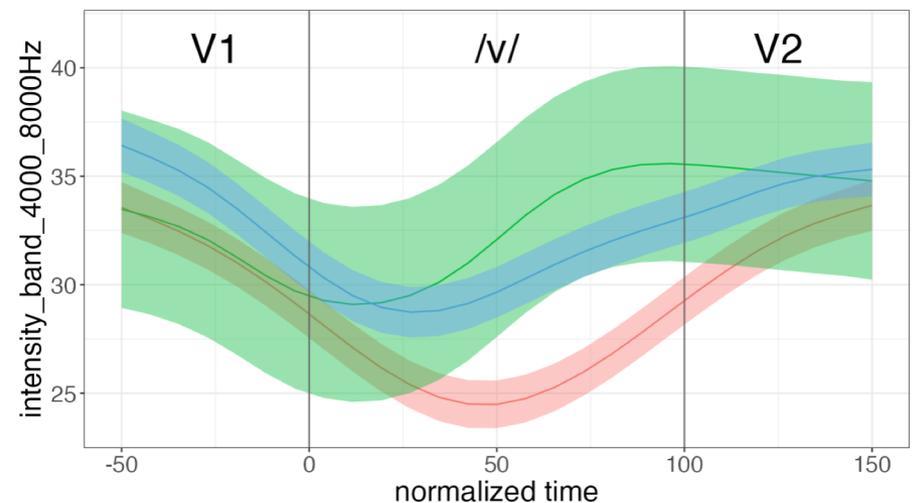


# Premiers résultats

- Mesures acoustiques qui discriminent bien les deux types de réalisation :
  - Intensité en basses et hautes fréquences



evaluation ■ approximante ■ fricative\_voisee ■ intermediaire



evaluation ■ approximante ■ fricative\_voisee ■ intermediaire

# DISCUSSION GENERALE

# Quelles mesures acoustiques choisir parmi le foisonnement de la littérature ?

- Possibilité d'extraire de nombreux paramètres acoustiques de façon automatisée à l'aide de scripts Praat, et plus encore d'outils tels qu'OpenSMILE
- Il peut être tentant de multiplier les paramètres, éventuellement en appliquant une méthode de réduction dimensionnelle ou de sélection de variable : dans ce cas, prudence !

# Quelles mesures acoustiques choisir parmi le foisonnement de la littérature ?

- Hormis des mesures très spécifiques qui nécessitent souvent une supervision fine pour s'assurer de leur fiabilité, la plupart des mesures acoustiques sont susceptibles de refléter de multiples phénomènes et pas uniquement celui ciblé
- C'est tout particulièrement le cas pour ceux liés au voisement et à la qualité de voix, comme HNR, CPP(S), etc.

# Phonétique de corpus ou approche expérimentale ?

- La plupart des études présentées aujourd'hui consistent en des études sur corpus
- Permet d'accéder à la parole telle qu'elle est produite dans un contexte plus naturel (réduction, disfluences, etc.), mais perte de contrôle expérimental
- La quantité de données et l'évolution des modèles stat. permet de compenser en partie, mais seulement jusqu'à un certain point

**Merci pour votre attention**

**Et merci à toutes les personnes qui  
ont contribué à ces études !**