

La focalisation prosodique en parole interprétée

RÉMI GODEMENT-BERLINE

LINGLUNCH, JEUDI 29 MARS 2018

Plan de la présentation

- Introduction
 - définitions
 - études précédentes
 - hypothèses
- Méthodologie
 - constitution et annotation du corpus
 - analyse prosodique
- Résultats
- Expérience de perception
 - méthodologie
 - résultats
- Conclusions

Introduction : définitions

Parole interprétée :

- oralisation d'un texte écrit préalablement mémorisé
- produit notamment par des comédien(ne)s
- diffère de la parole lue et de la parole spontanée

Focalisation prosodique :

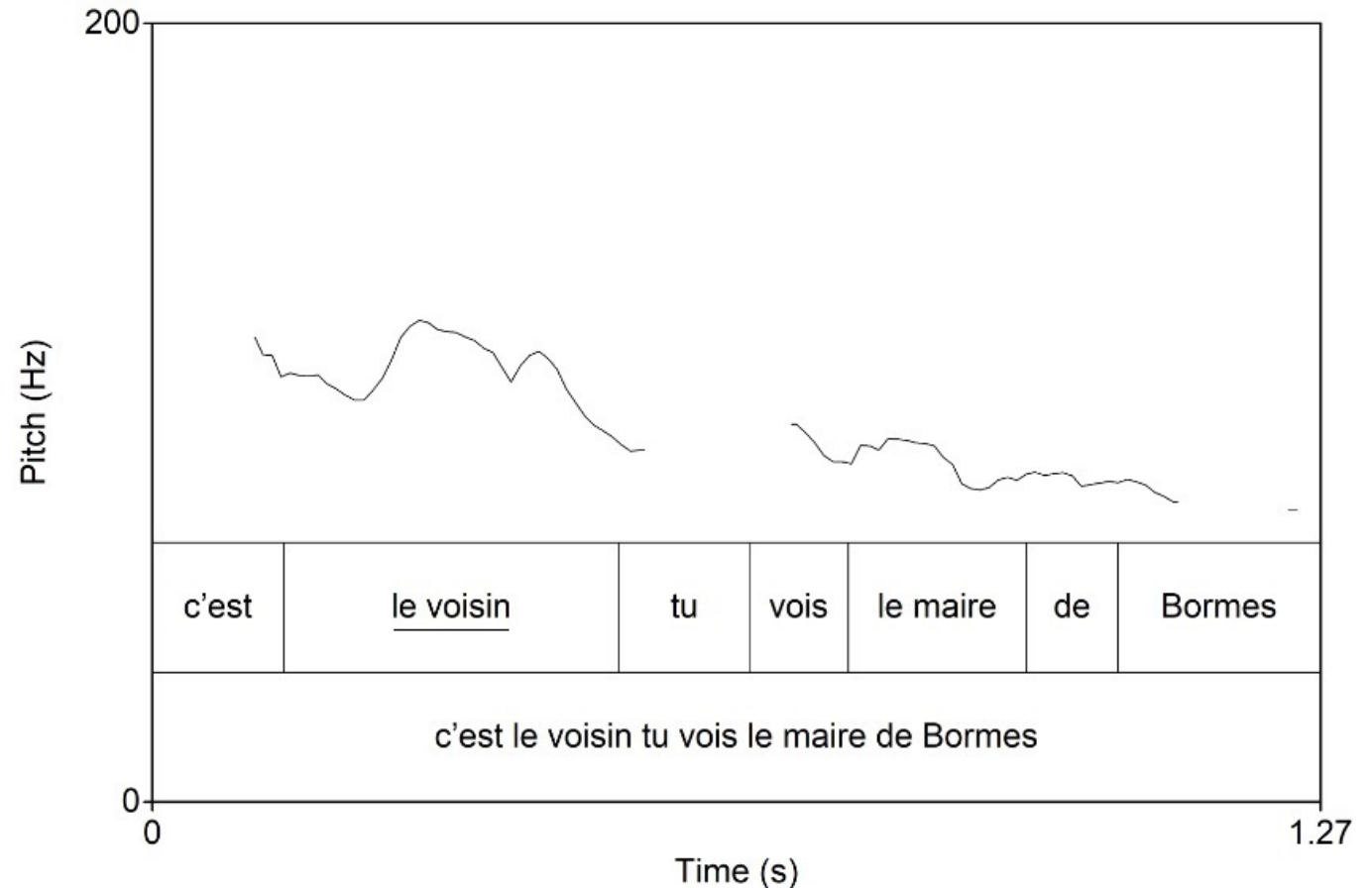
- soulignement d'un constituant dans un énoncé
- différentes ressources prosodiques
- remplit 2 ensembles de fonctions :
 - marquage de focus
 - emphase

Introduction : traits de la focalisation

Principaux traits prosodiques :

- augmentation de hauteur, de durée et d'intensité
- accent initial
- contour terminal
- désaccentuation post-focale

Di Cristo 1999, Rossi 1999, Jun et Fougeron 2000, Astésano 2001, Post 2002, Martin 2009, Delais-Roussarie et al 2015



Introduction : phonogenres

→ différences de **conditions de production** (Goldman, Auchlin et Simon 2011)

	parole spontanée	parole lue	parole interprétée
caractère médiatique	non-médiatique	non-médiatique	médiatique ou non-médiatique
type d'audience	face à face	micro	public ou caméra
degré de préparation	improvisé	lu à haute voix	mémorisé

Introduction : différences prosodiques entre phonogenres

Parole spontanée/lue :

- beaucoup de résultats contradictoires
- parole lue :
 - plus de déclinaison de la hauteur
 - plus de contours descendants
 - plus de syllabes accentuées
 - groupes prosodiques plus courts
 - allongements de fin de groupes plus marqués
 - beaucoup moins de marques d'hésitation
 - moins de perturbation de la qualité vocalique

Lieberman et al 1985, Howell et Kadi-Hanifi 1991, Blaauw 1994, Batliner et al 1995, Swerts et al 1996, Laan 1997, Astésano 2001, Mixdorff et Pfitzinger 2005, Simon et al 2010, Goldman et al 2011, de Ruiter 2015, Dellwo et al 2015, Wagner et Windmann 2016

Introduction : différences prosodiques entre phonogenres

Parole spontanée/interprétée :

- plusieurs études sur d'autres aspects :
 - capacité de discrimination auditive
 - jugements d' « intensité émotionnelle »
 - différences entre acteurs professionnels et amateurs
- parole interprétée :
 - registre tonal plus élevé Audibert et al 2010
 - plus de variation de la hauteur Jürgens et al 2011
 - différences de qualité vocalique

Audibert et al 2008, Jürgens et al 2011

Audibert et al 2008

Jürgens et al 2015, Kraemer et Swerts 2008

Introduction : hypothèses

Question de recherche principale :

La focalisation prosodique présente-t-elle des propriétés spécifiques en parole interprétée ?

Hypothèse 1 :

La fréquence d'occurrence de la focalisation augmente en parole interprétée

Hypothèse 2 :

La focalisation est plus marquée prosodiquement en parole interprétée

Hypothèse 3 :

Les fonctions de la focalisation présentent des réalisations distinctes.

Méthodologie : constitution du corpus

Protocole RepTask (cf. Laurens, Marandin, Patin et Yoo 2011) :

- enregistrement de 2 dialogues semi-spontanés
- reproduction par 4 locuteurs en parole lue et interprétée

Donc j'étais à ce cours-là, et c'était assez drôle parce que tous les étudiants se sont précipités au cours de philosophie indienne parce que c'était plutôt original... Et le prof était tellement horrible...

spontané :



lecture :



interprétation :



Méthodologie : annotation du corpus

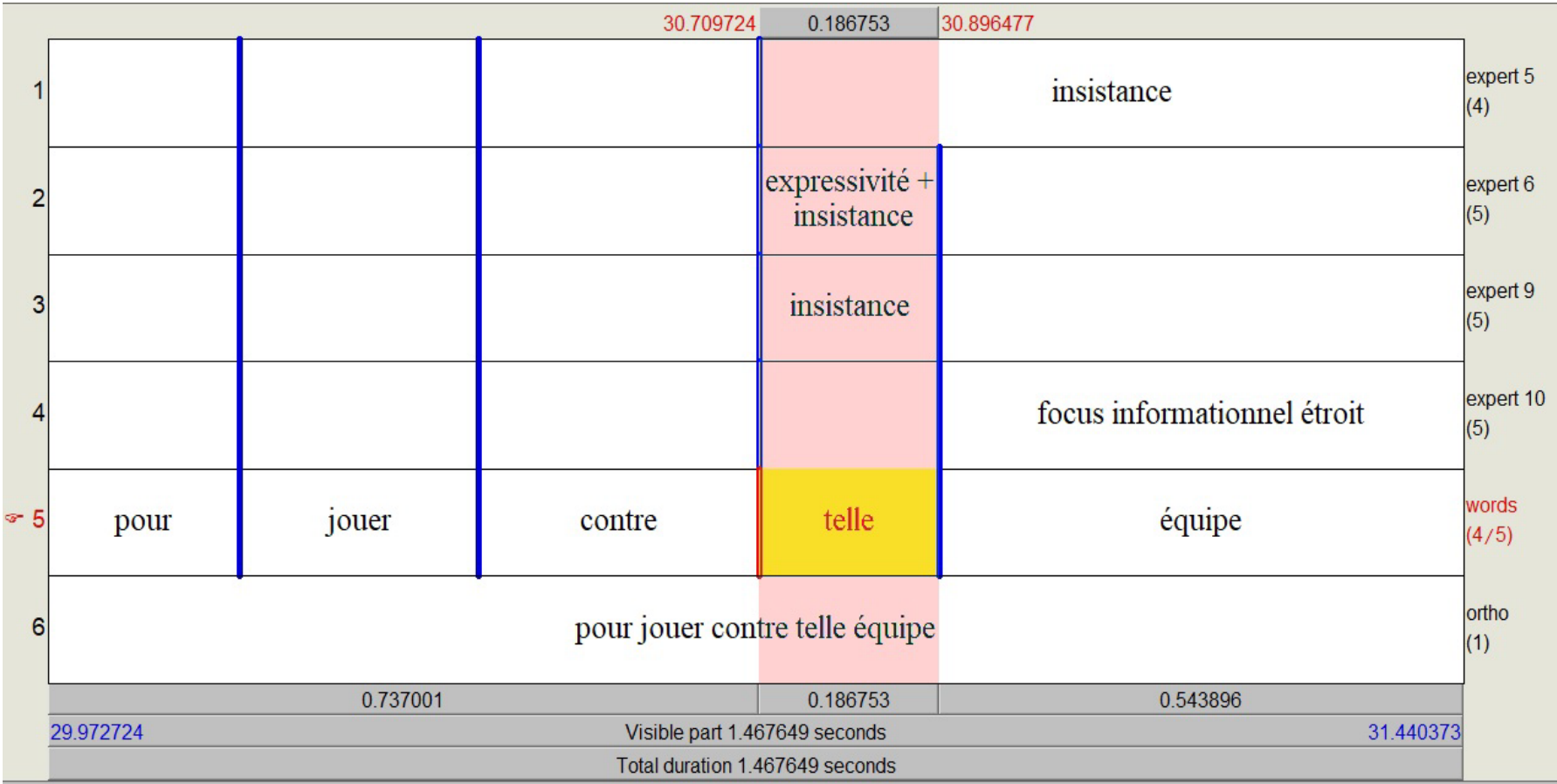
Relevé des occurrences de focalisation :

- groupes de 4 experts par enregistrement
- relevé auditif uniquement
- critères acoustiques précis

Classification fonctionnelle des occurrences :

- mêmes groupes d'experts
- choix parmi plusieurs fonctions

Méthodologie : annotation du corpus



Méthodologie : analyse prosodique

Segmentation semi-automatique du corpus en mots, syllabes et phones (EasyAlign)

Mesures automatiques de hauteur et de durée sur chaque syllabe (Prosogram)

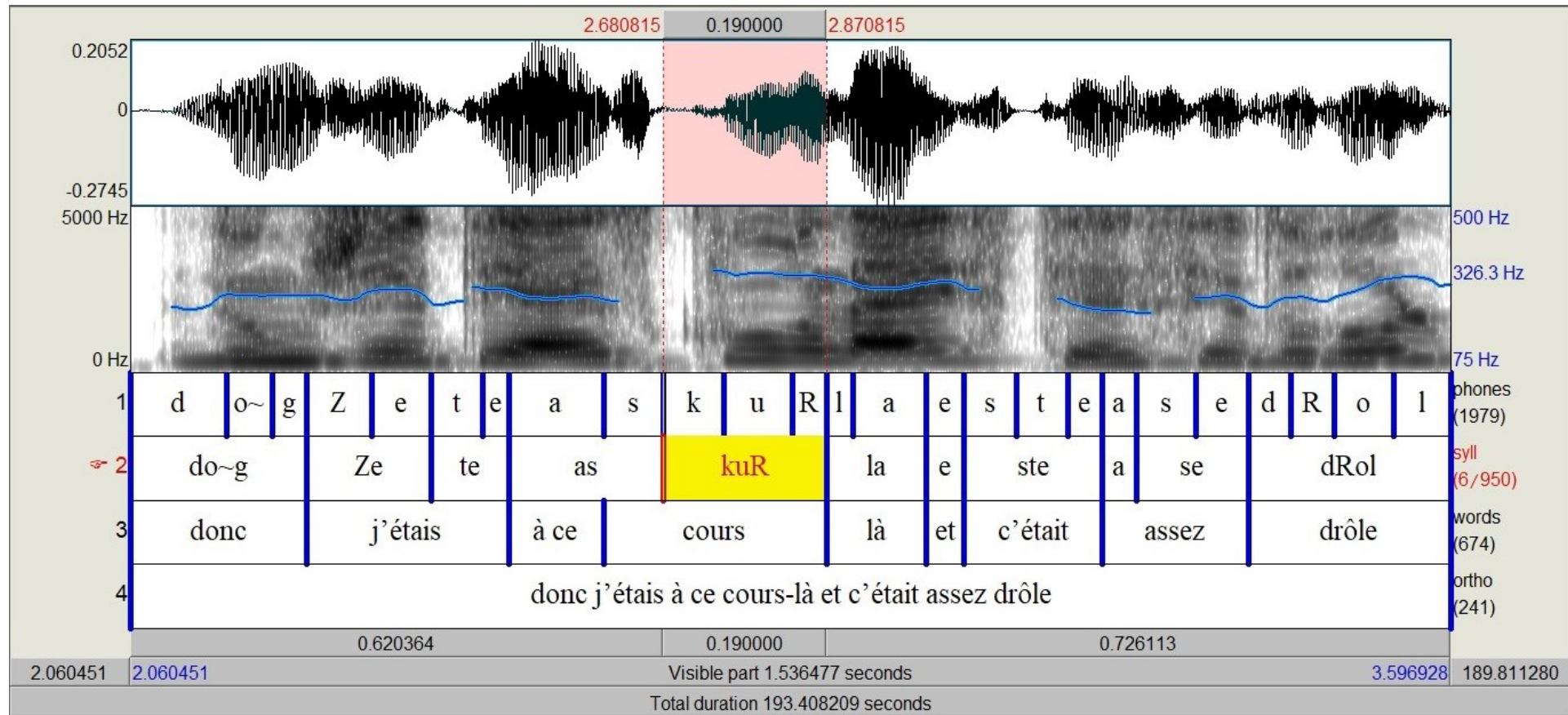
- hauteur normalisée par rapport aux locuteurs
- durée normalisée par rapport à la structure syllabique et au débit de parole des locuteurs

Description phonologique des constituants focalisés selon 3 catégories :

- contour intonatif : transcription manuelle en F-ToBI (cf. [Delais-Roussarie et al 2015](#))
- étendue syllabique du contour intonatif
- présence d'accent initial : détection automatique des syllabes proéminentes (Analog)

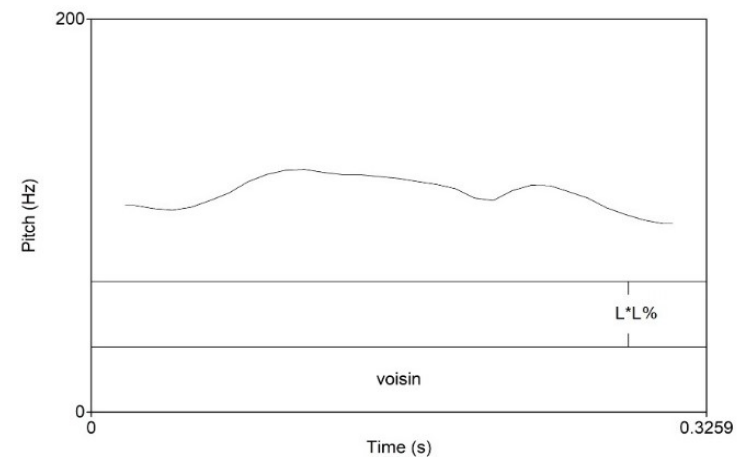
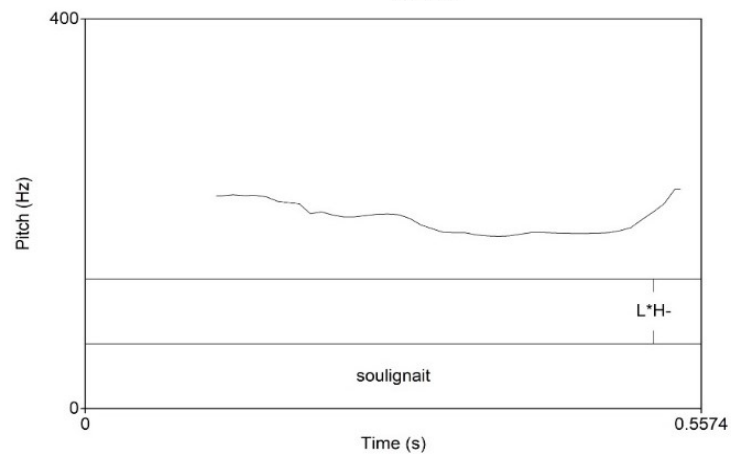
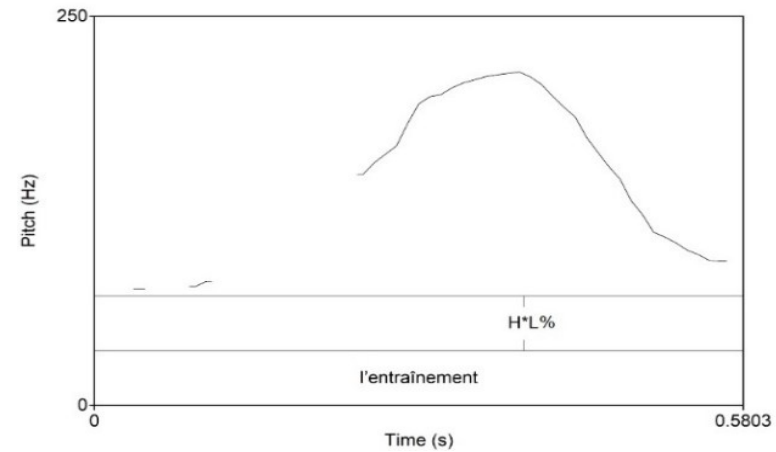
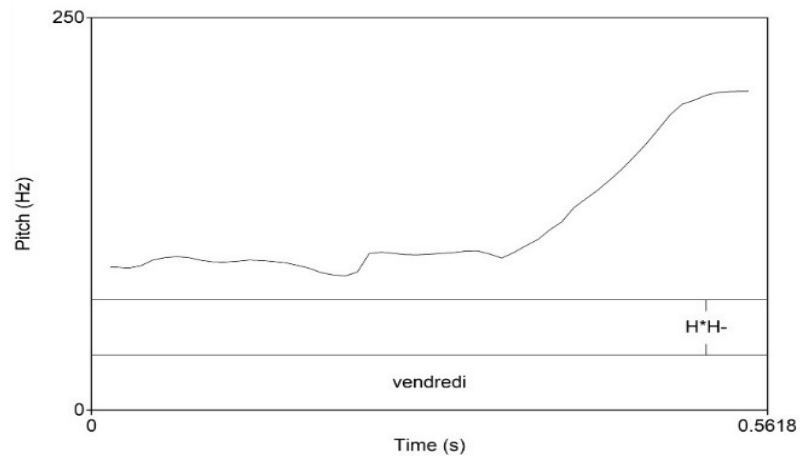
Méthodologie : analyse prosodique

Segmentation du corpus (EasyAlign)



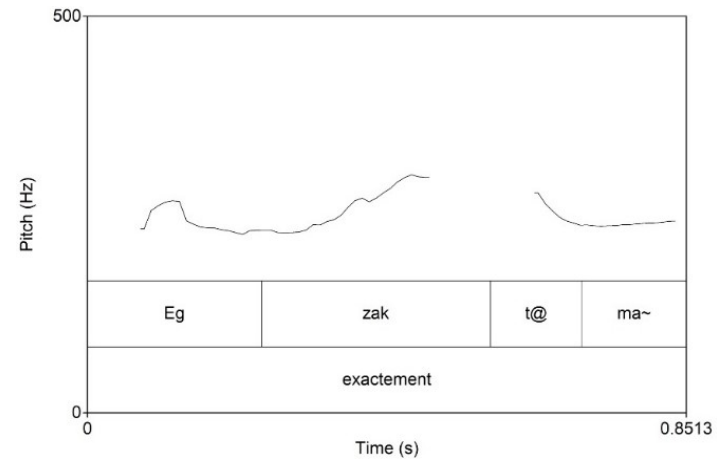
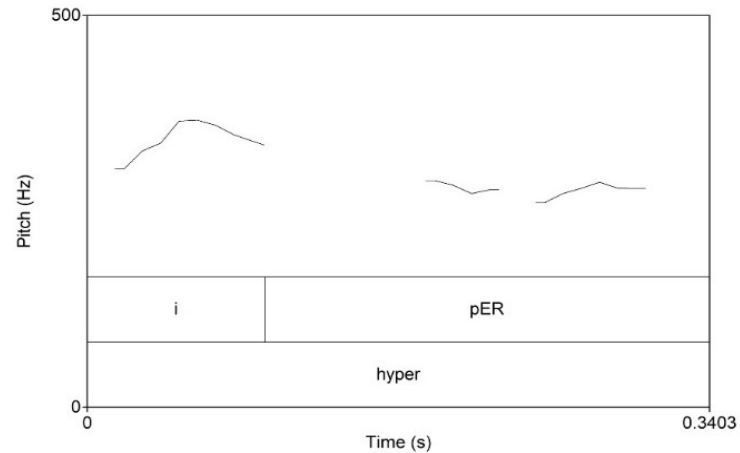
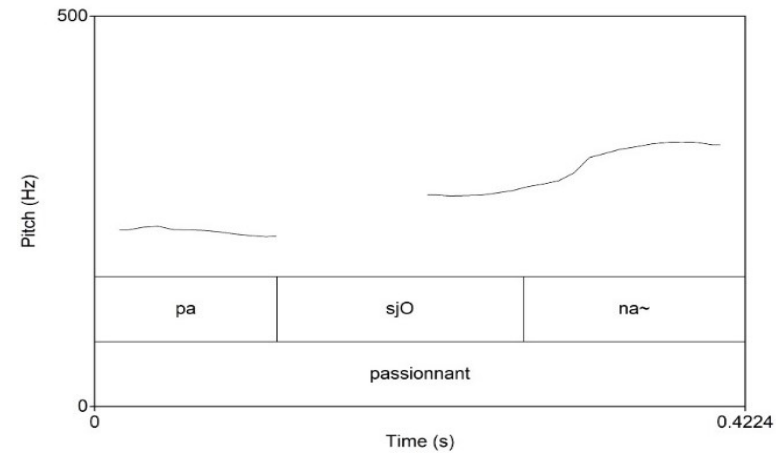
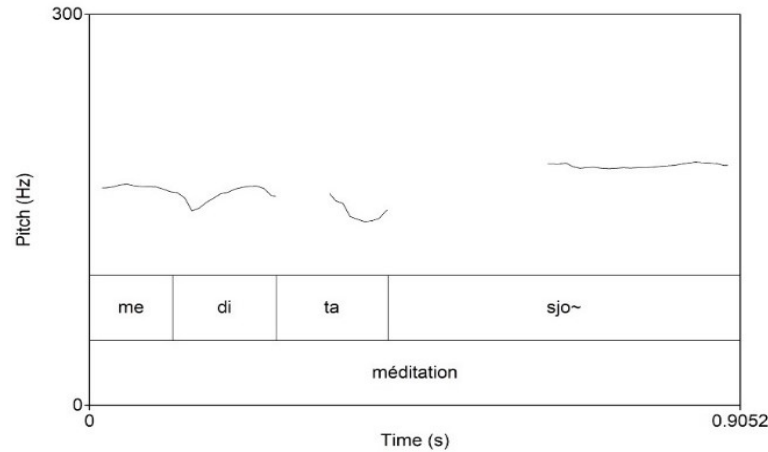
Méthodologie : analyse prosodique

Contours intonatifs (F-ToBI)



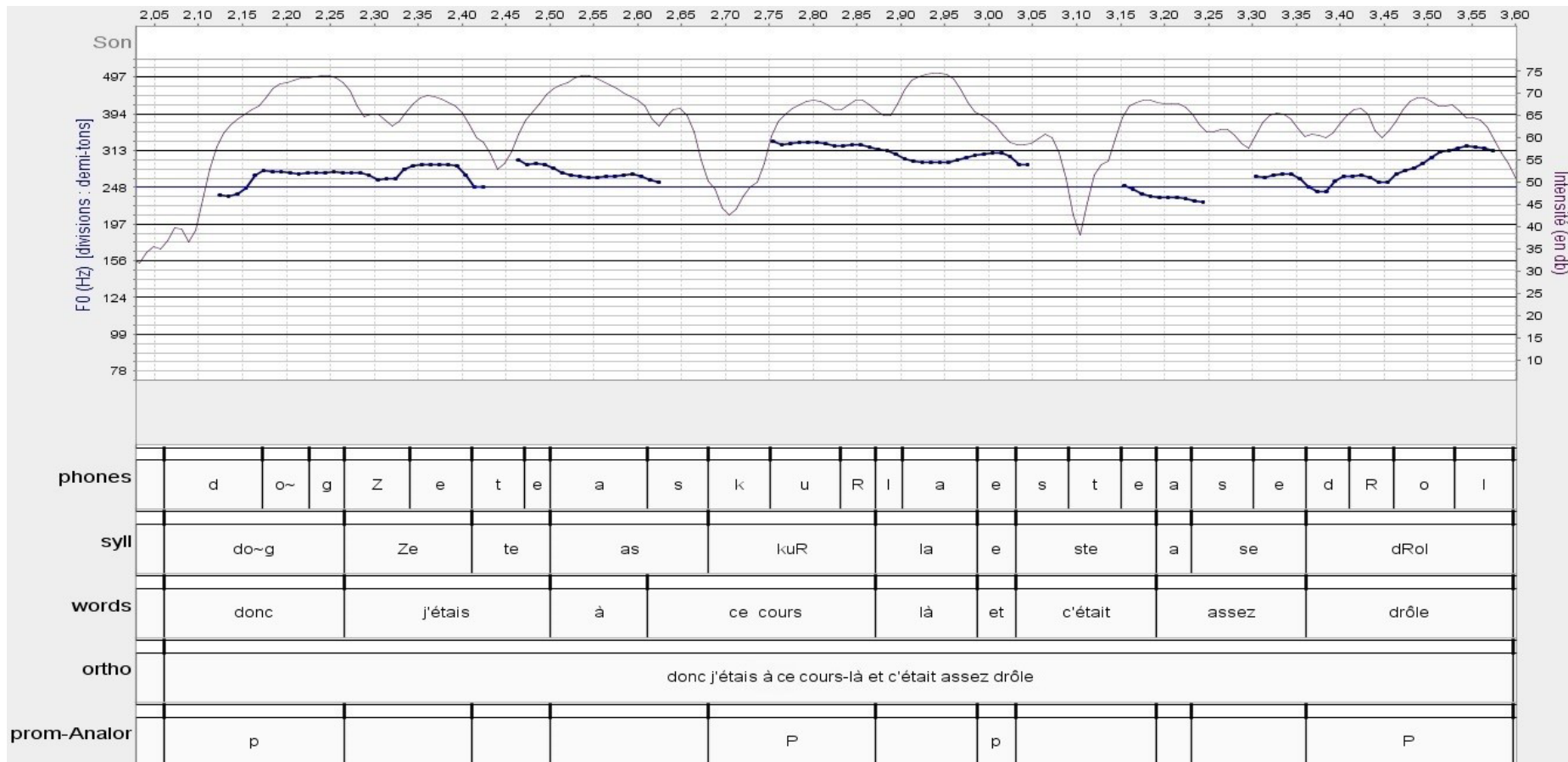
Méthodologie : analyse prosodique

Etendue syllabique des contours



Méthodologie : analyse prosodique

Détection d'accent initial (Analog)



Taux d'accord entre les experts

Relevé des occurrences de focalisation :

	parole spontanée	parole lue	parole interprétée
kappa de Fleiss	0,374	0,230	0,285

Classification fonctionnelle des occurrences :

	marquage de focus	insistance	expressivité
kappa de Fleiss	0,121	0,166	0,183

Plusieurs facteurs possibles :

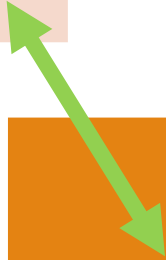
- difficulté d'effectuer un relevé purement auditif
- disparité dans la compétence des experts
- présence d'un phonostyle du comédien
- différence de naturalité entre les phonogenres
- répartition des enregistrements parmi les experts
- définitions des fonctions mal assimilées/respectées
- pas assez d'occurrences pour certaines fonctions
- pas de distinction prosodique entre les fonctions

Fréquence d'occurrence

	parole spontanée	parole lue	parole interprétée	corpus entier
pourcentage de syllabes focalisées	6,6 %	10,7 %	14,2 %	11,3 %

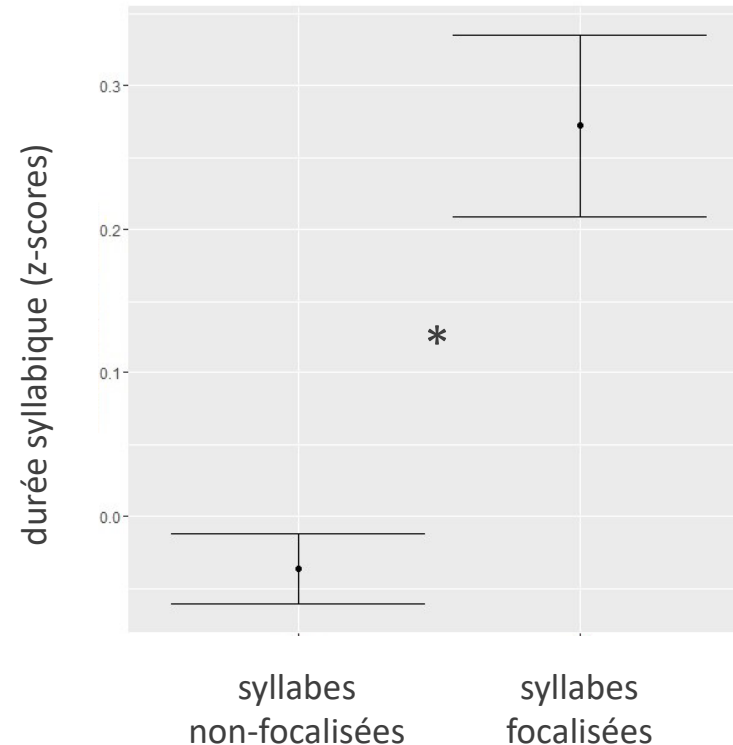
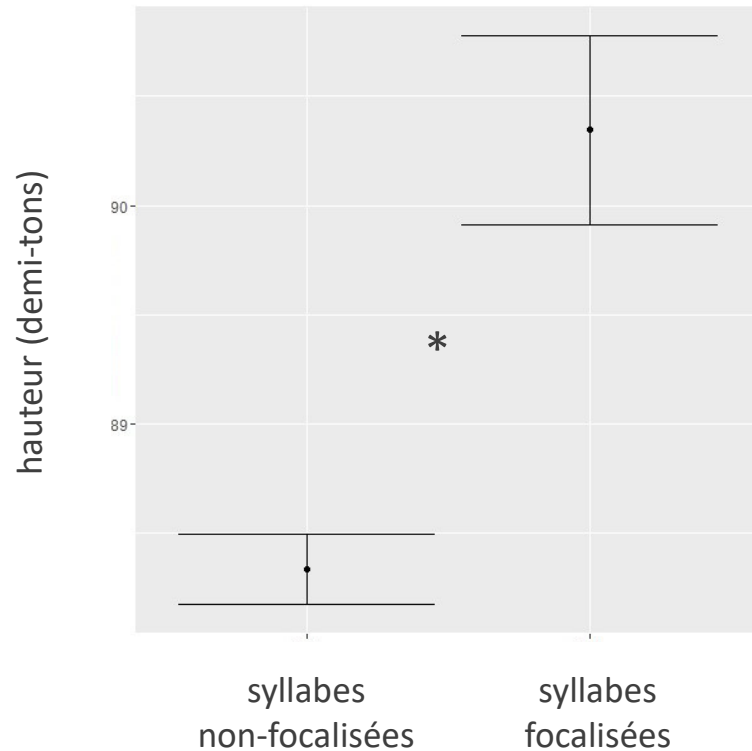
effet du phonographe significatif :
 $\chi^2(2) = 22,7, p < 0,001$

- hypothèse 1 validée
- relation avec le degré de « préparation » (cf. Goldman, Auchlin et Simon 2011)
- méthodologie : parole interprétée = phonographe favorable à l'étude de la focalisation



	corpus entier
pourcentage de syllabes proéminentes	35,8 %

Focalisation : hauteur et durée syllabique



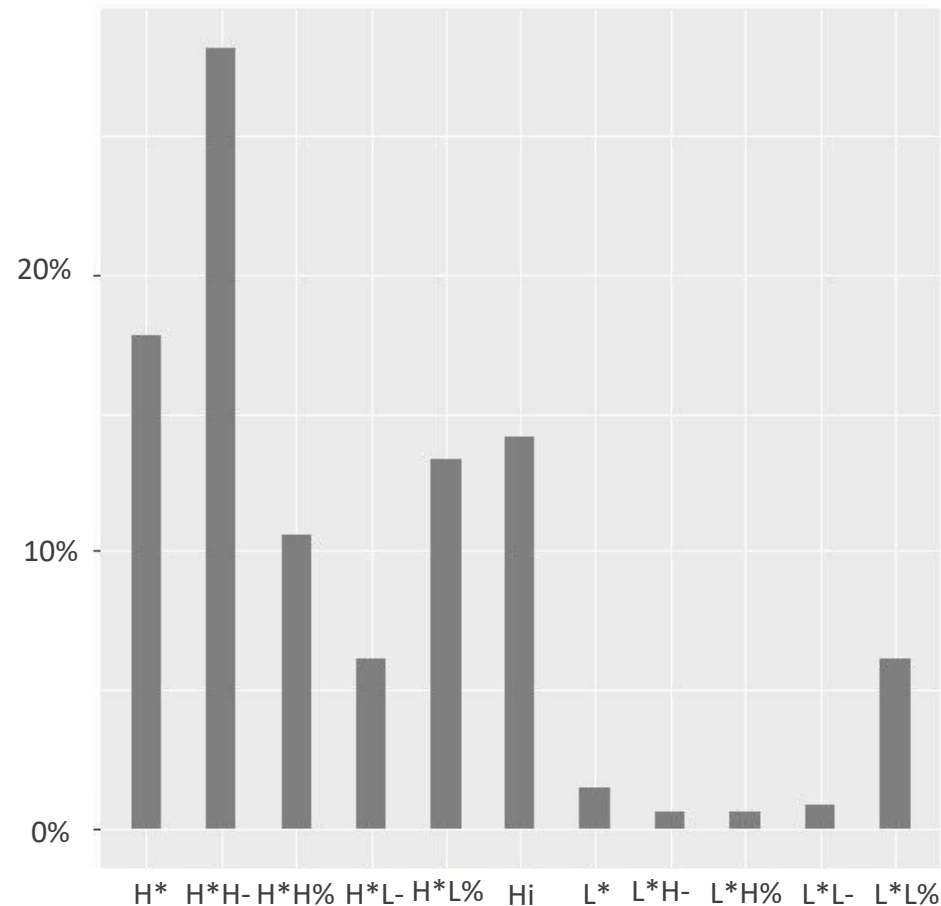
effet de la présence de focalisation significatif :

hauteur : $\chi^2(1) = 521,3$, $p < 0,001$

durée : $\chi^2(1) = 71,8$, $p < 0,001$

➤ confirme plusieurs études précédentes (e.g. Di Cristo 1999b, Rossi 1999, Astésano 2001)

Focalisation : contours intonatifs



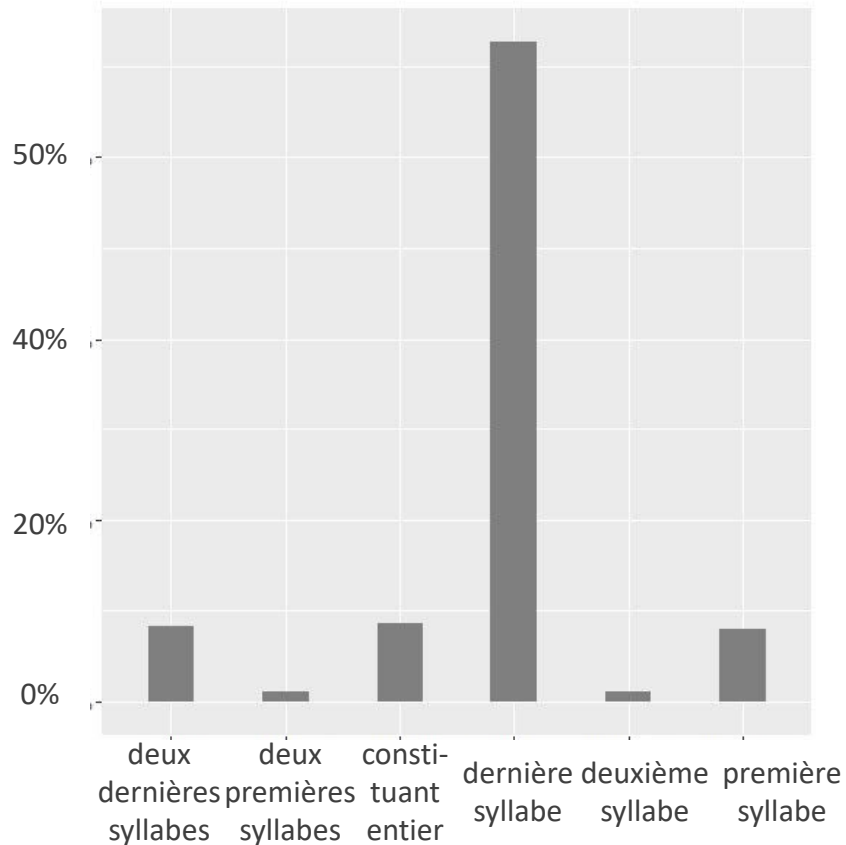
Contours intonatifs :

- commencent majoritairement par un ton haut
- majorité de contours montants
- contour le plus fréquent : H*H-

Frontières prosodiques :

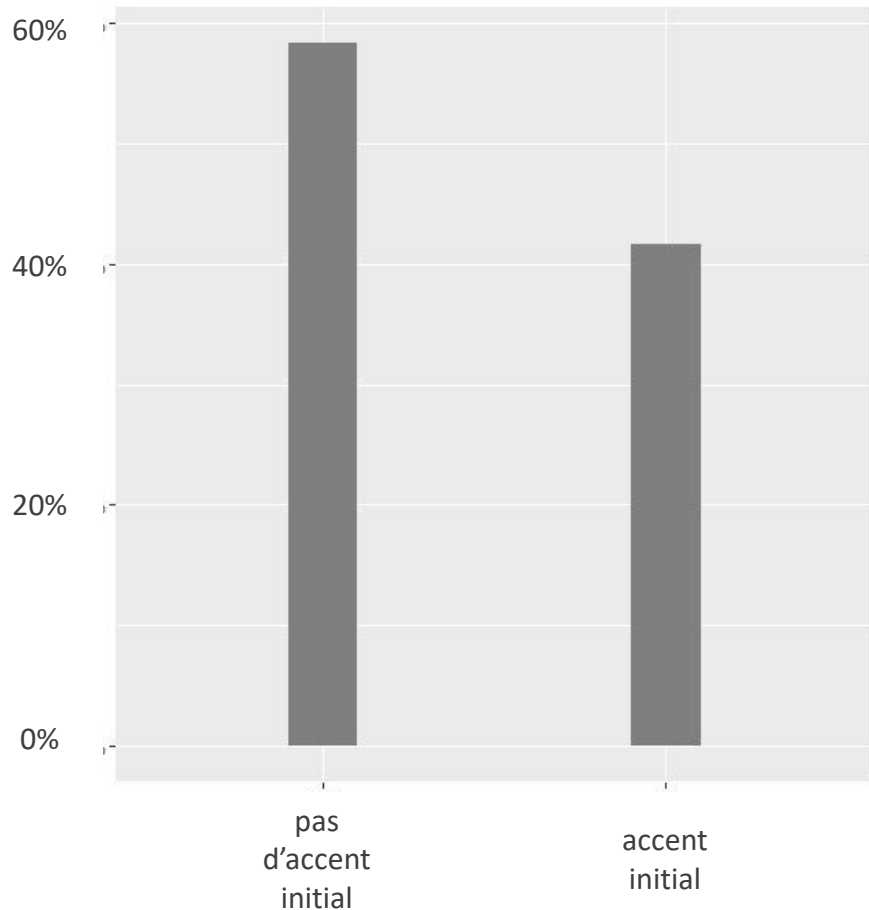
- la focalisation a lieu à tous les niveaux de la structure prosodique
- cohérent avec les différentes fonctions attribuées à la focalisation

Focalisation : étendue du contour



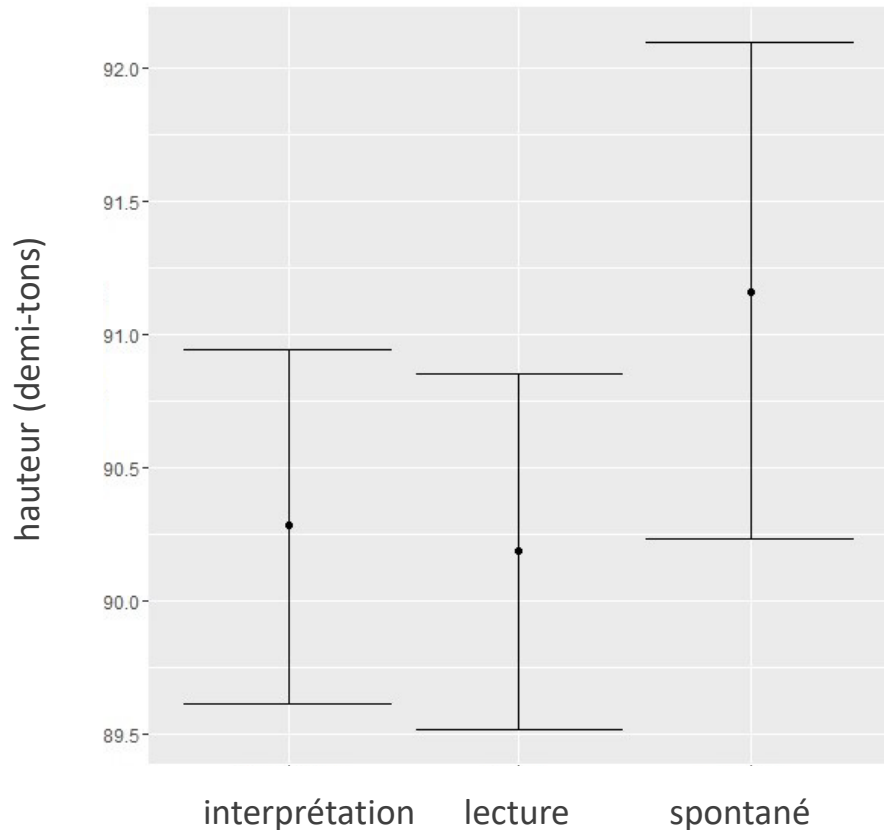
- contours portés en majorité par la dernière syllabe
- distribution par rapport au type de contour :
 - dernière syllabe : contours montants ou descendants
 - deux dernières syllabes ou constituant entier : contours bitonaux montants-descendants ou descendants-montants
 - première syllabe ou deux premières syllabes : ton haut initial Hi

Focalisation : accent initial



- trait fréquent mais pas obligatoire
- confirme plusieurs études précédentes (e.g. Di Cristo 1999, Rossi 1999, Astésano 2001)

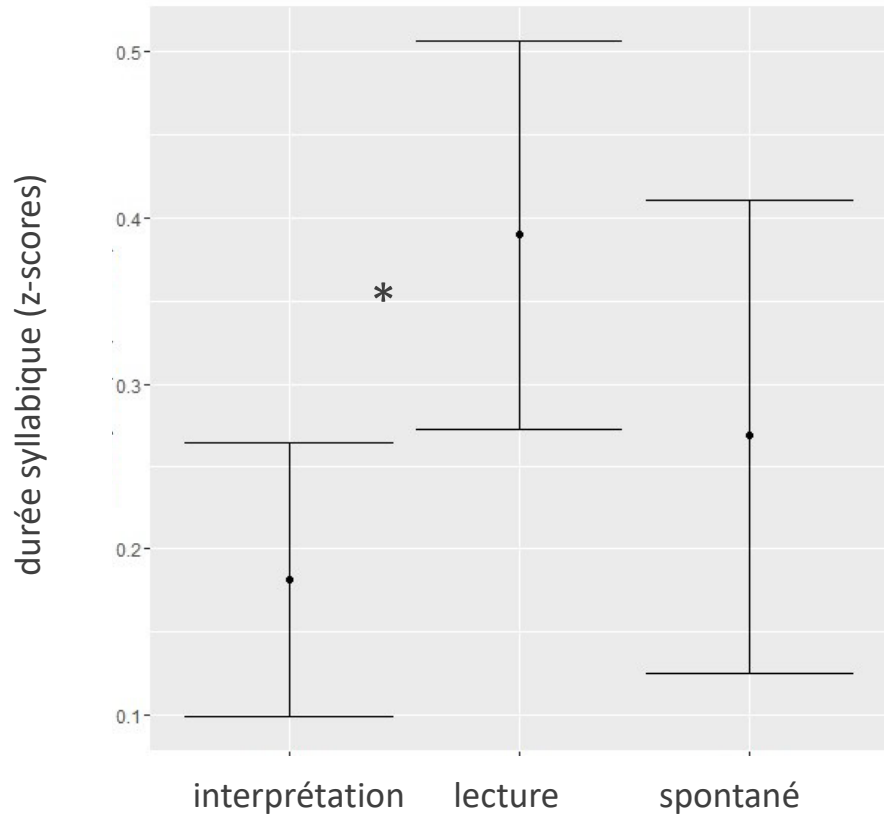
Différences entre phonogenres : hauteur



effet du phonogenre non significatif :
 $\chi^2(2) = 0,3, p > 0,1$

- aucune différence significative de hauteur
- parole interprétée/spontanée : contradictoire avec les études précédentes
- parole lue/spontanée : pas forcément surprenant car la littérature est contradictoire
- facteur possible : effet inhibiteur du contexte expérimental sur les locuteurs de parole interprétée

Différences entre phonogenres : durée syllabique



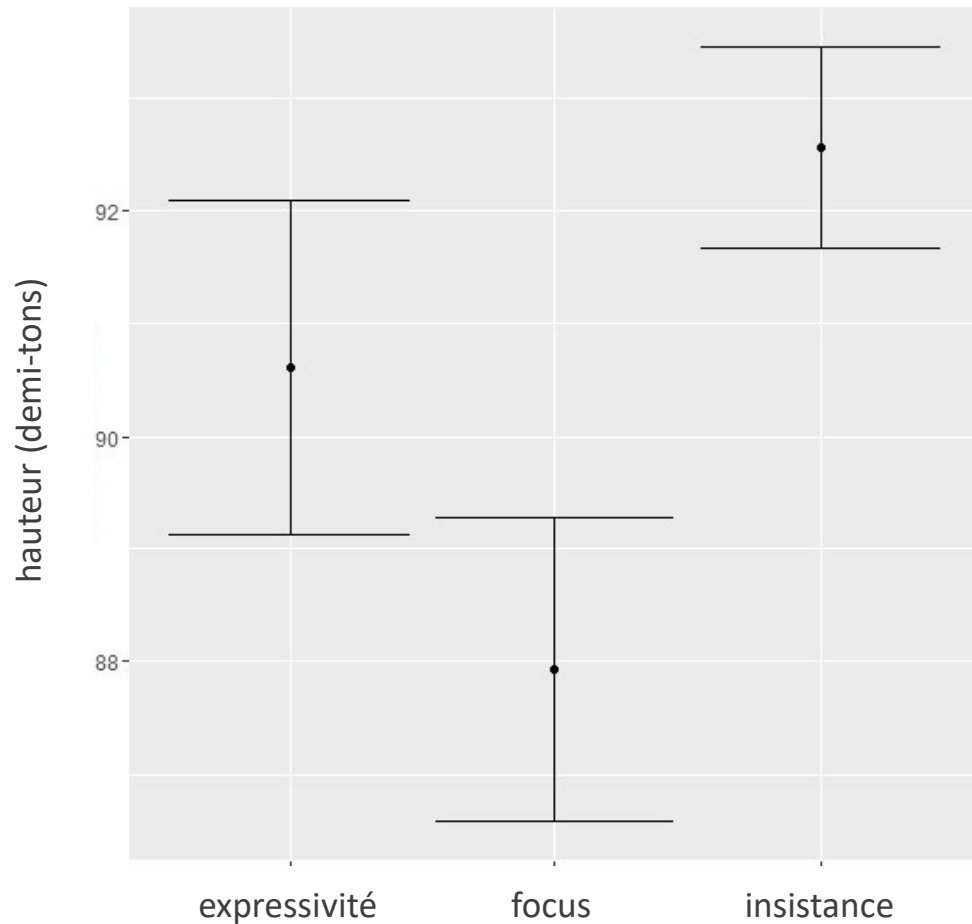
effet du phonogène significatif :
 $\chi^2(2) = 9,1, p < 0,05$

➤ parole lue : trait caractéristique ?

Différences entre phonogenres : catégories phonologiques

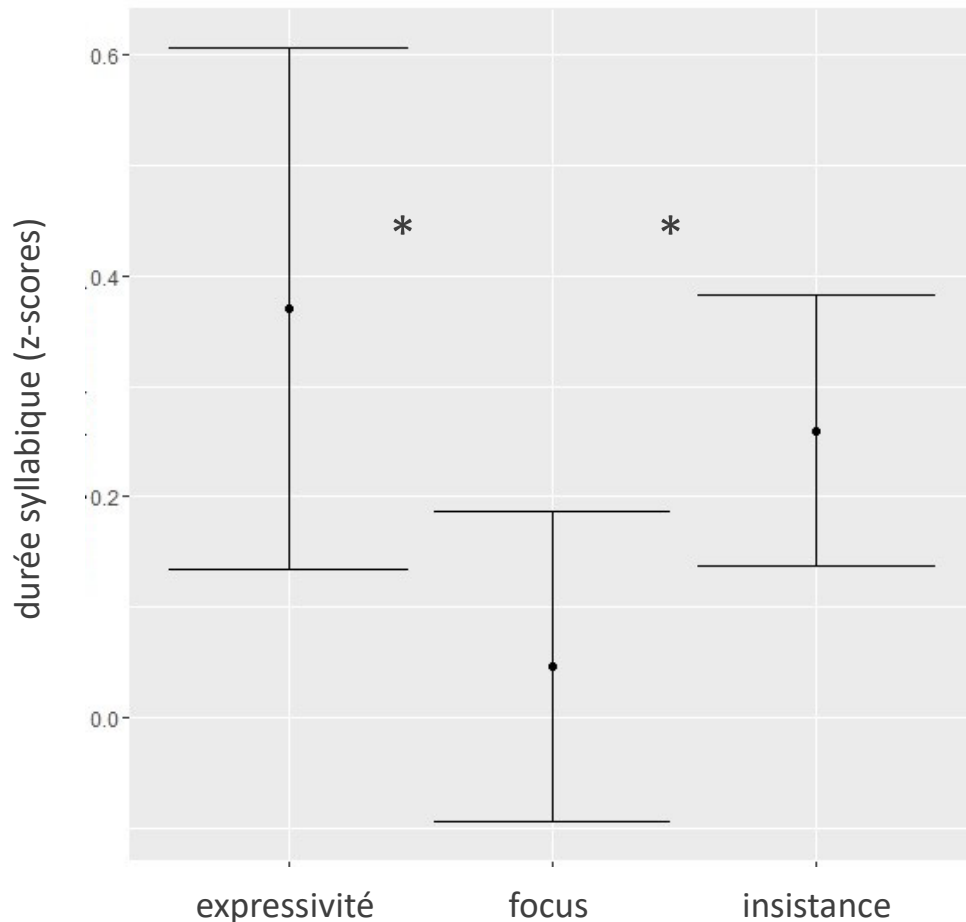
- **aucune différence significative**
- contour intonatif : surprenant étant donné la grande variété observée
- étendue du contour : moins surprenant car les contours sont majoritairement portés par la dernière syllabe
- accent initial : contredit Goldman et al 2011

Différences entre fonctions : hauteur



effet de la fonction non significatif :
 $\chi^2(2) = 4,2, p > 0,1$

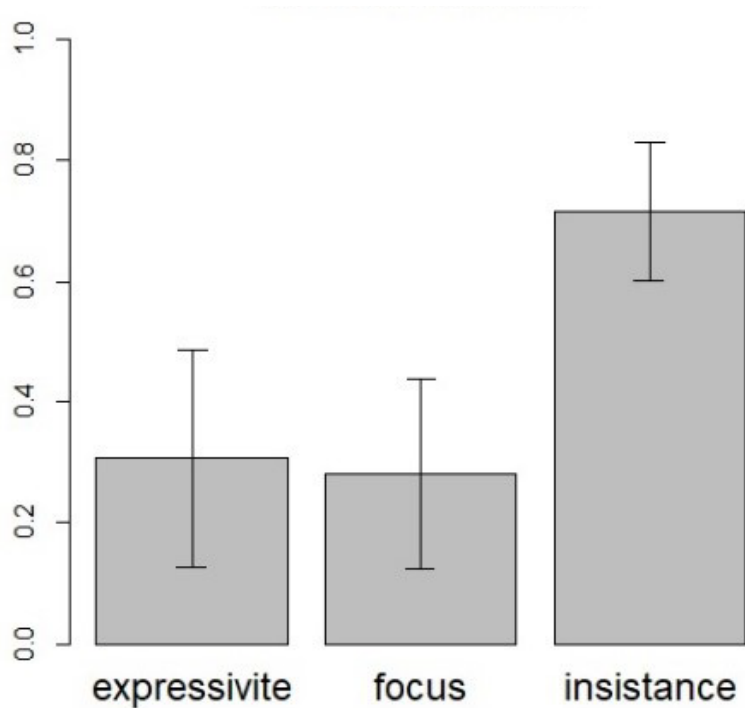
Différences entre fonctions : durée



effet de la fonction significatif :
 $\chi^2(2) = 6,9, p < 0,05$

- suggère qu'il existe bien une distinction phonétique entre les fonctions emphatiques et les fonctions sémantico-pragmatiques de la focalisation
- expressivité : cohérent avec la littérature

Différences entre fonctions : accent initial



effet de la fonction significatif :
 $\chi^2(2) = 21,8, p < 0,001$

- confirme la relation, souvent décrite dans la littérature, entre insistance et accentuation initiale

Expérience de perception

Objectif : vérifier les résultats obtenus en production concernant les différences entre fonctions

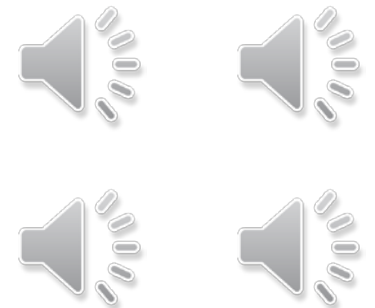
- durée syllabique : focus < insistance et expressivité
- relation entre insistance et accent initial

Protocole expérimental :

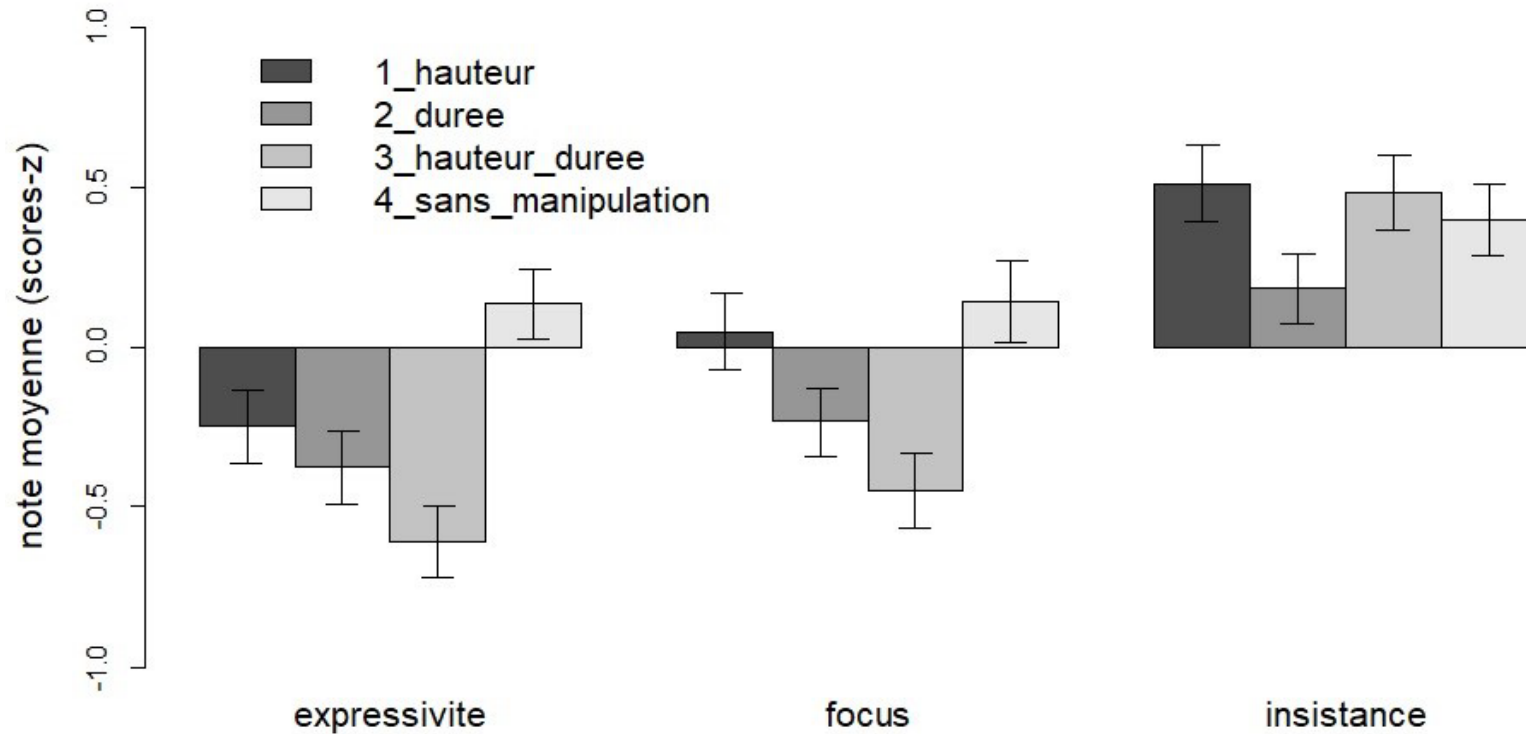
- obtention de jugements d'acceptabilité
- 30 participants
- 30 énoncés extraits du corpus contenant une focalisation (10 par fonction)
- manipulation prosodique de la 1^{ère} syllabe du constituant focalisé (Momet-Intsint) :
 - augmentation de hauteur
 - augmentation de durée
 - augmentation de hauteur et de durée
 - pas de manipulation
- présentation visuelle et auditive de l'énoncé et de son contexte d'origine
- évaluation de l'acceptabilité de chaque réalisation par rapport au contexte

contexte : *A propos de mon nom, du coup j'ai une anecdote. Quand j'étais... Je devais être...*

énoncé : *Ouais, j'étais au LYCEE.*



Réalisation des fonctions



effet de la fonction significatif :
 $\chi^2(2) = 18,3, p < 0,001$

- Le fait que l'insistance soit globalement mieux évaluée que les autres fonctions confirme indirectement la relation entre insistance et accentuation initiale, déjà observée en production.

Conclusion : hypothèses

Hypothèse 1 : la fréquence d'occurrence de la focalisation augmente en parole interprétée

→ confirmée

Hypothèse 2 : la focalisation est plus marquée prosodiquement en parole interprétée

→ non confirmée

Hypothèse 3 : les fonctions de la focalisation présentent des réalisations distinctes

→ en partie confirmée

Conclusion : résumé

Principaux apports :

- analyse de l'influence du phonogène sur la fréquence d'occurrence et la réalisation de la focalisation
- protocoles expérimentaux relativement novateurs
- comparaison de la réalisation prosodique de l'insistance, l'expressivité et le marquage de focus
- confirmation de certains des traits attribués à la focalisation
- évaluation de la fréquence d'occurrence de la focalisation en français

Conclusion : perspectives

Consolidation de la méthodologie :

- obtenir plus de données afin de pouvoir analyser davantage de fonctions de la focalisation
- recours à des participants lambda (ex : *crowdsourcing*) portant des jugements naïfs
- perception : manque de naturalité des stimuli

Prise en compte de traits prosodiques supplémentaires :

- intensité, qualité vocalique
- configuration tonale et constituance syllabique des accents de focalisation (cf. *Astésano 2001*)
- désaccentuation pré-focale ou post-focale