

IDENTIFICATION DU PROCESSUS D'INNOVATION PHONETIQUE

Etat des lieux d'un projet

Florent Chevalier
Université de Poitiers



“Des savoirs & des talents”

Réunion d'équipe
FoReLLIS A,
5 avril 2019

I – Présentation du projet

II – Avancement

III – Difficultés rencontrées

IV – Etapes suivantes



I – PRESENTATION DU PROJET

Cadre institutionnel, contexte
théorique, méthodologie et
objectifs



I – PRESENTATION DU PROJET : CADRE INSTITUTIONNEL

Thèse de doctorat :

« De la convergence inter-locuteurs au changement phonétique : accommodation linguistique et changement phonétique dans l'anglais parlé à Glasgow. »

Financée par un contrat doctoral ministériel 2017-2020

Sous la direction de Sylvie Hanote (FoReLLIS, UP) et Jane Stuart-Smith (GULP, University of Glasgow)



I – PRESENTATION DU PROJET : CONTEXTE THEORIQUE

- Le changement phonétique et ses étapes
- Où/quand/pourquoi a lieu l'innovation ?
- Hypothèse de Trudgill (1986) : l'innovation pourrait correspondre au phénomène d'accommodation linguistique à très court terme



I – PRESENTATION DU PROJET : CONTEXTE THEORIQUE

- L'accommodation : un concept de psychologie sociale (Giles & St Clair 1979). Observée au niveau non verbal, paraverbal, et verbal.
- Lien entre accommodation et changement phonétique étudié à court terme (quelques minutes, Babel 2009) ainsi qu'à moyen terme (après plusieurs semaines, Sonderegger *et al.* 2017)
- Absence d'étude à long terme (plusieurs décennies)



I – PRESENTATION DU PROJET : CONTEXTE THEORIQUE

- Soulève la question des méthodes utilisées en sociophonétique
- On part en général du principe que les locuteurs ont une prononciation absolument stable, et on mesure le changement à l'aide de moyennes (Cukor-Avila & Bailey 2013)
- Méthodes vouées à ne pas repérer l'innovation ?

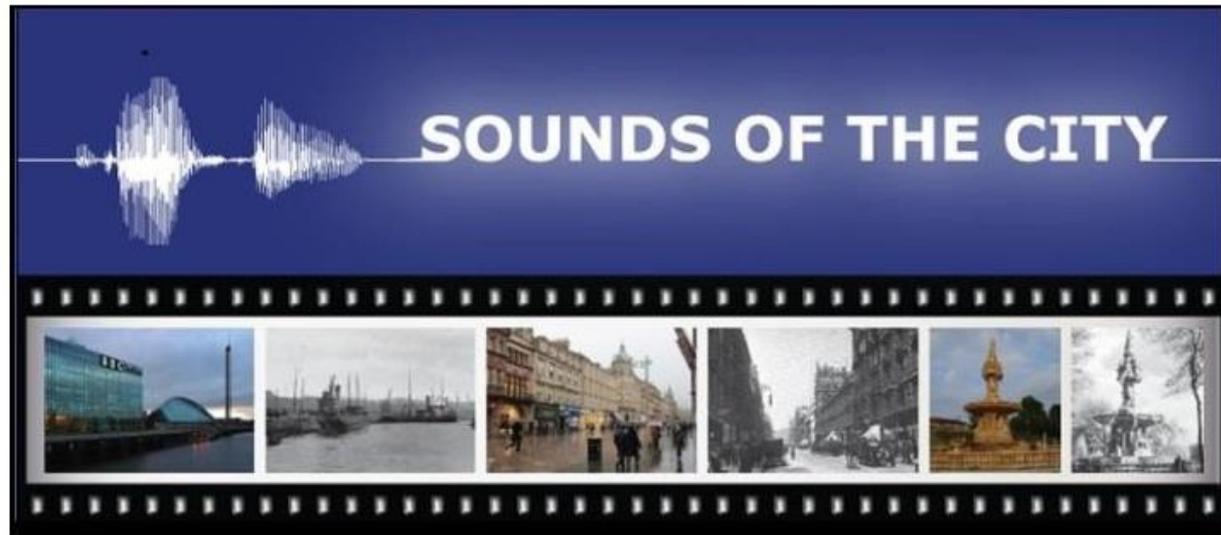


Corpus idéal pour étudier la relation entre convergence et changement phonétique (Pardo 2013) :

- production orale spontanée
- différents types d'interaction
- locuteurs d'âge et de genre variés
- dans une communauté linguistique fermée
- offrant une perspective en diachronie



I – PRESENTATION DU PROJET : METHODOLOGIE



Decade of Recording	Old 67-90 (Decade of Birth)	Middle-aged 40-55 (Decade of Birth)	Young 10-17 (Decade of Birth)
1970s	4 f, 6 m (1890s)	7 f, 7 m (1920s)	4 f, 8 m (1960s)
1980s	6 f, 6 m (1900s)	4 f, 12 m (1930s)	2 f, 5 m (1970s)
1990s	6 f, 6 m (1910s)	6 f, 6 m (1940s)	6 f, 6 m (1980s)
2000s	6 f, 6 m (1920s)	6 f, 5 m (1950s)	6 f, 6 m (1990s)

Table 1: Real- and apparent-time structure of the main Glasgow corpus; an additional 6 speakers born in the 1890s and recorded in 1916/17 are also available. Recordings are predominantly of unprepared spontaneous speech from sociolinguistic surveys, oral histories and broadcast media.

I – PRESENTATION DU PROJET : METHODOLOGIE

- Etude de la réalisation des segments vocaliques (caractéristiques temporelles et spectrales)
- Analyse de la variation dynamique (à l'intérieur de l'interaction), ainsi qu'en diachronie large (temps réel et apparent)
- Comparaison des modalités de la variation



I – PRESENTATION DU PROJET : OBJECTIFS FINAUX

- Contribuer à la recherche fondamentale en socio-linguistique en explicitant le lien entre accommodation et changement phonétique
- Identifier les facteurs linguistiques, situationnels et sociaux susceptibles d'encourager ou de freiner la convergence entre locuteurs
- Contribuer à la description de l'anglais de Glasgow et de son évolution à travers le XXème siècle



II – AVANCEMENT DU PROJET

Le premier pilote : ses raisons
d'être, ses tentatives, et ses
premiers résultats



Pourquoi faire un premier pilote ?

- Définir une méthode de travail systématique : étapes à effectuer, dans quel ordre, avec quels critères ...
- Evaluer les valeurs extraites du corpus, obtenues au moyen d'une segmentation automatique (HTK)
- Me former aux outils nécessaires



II – AVANCEMENT DU PROJET : PREMIER PILOTE

Sous-corpus utilisé pour le pilote :

- Toutes les interactions entre pairs (19) :

	90M	90Y	00M	00Y
F	2	2	3	3
M	2	2	2	3

- Segments retenus : /i ʌ ɔ o/
(mots-clés : *FLEECE*, *FOOT/GOOSE*, *LOT*, *GOAT*)



Préparation des données

- Extractions de toutes les occurrences de /i ɥ ɔ o/ sous accent lexical via l'interface LaBB-CAT
- Retrait des incongruités
- Annotation selon le lieu d'articulation des segments précédent et suivant



Préparation des données

- Extraction des valeurs spectrales (F1 et F2 pris aux instants 0, 0.5 et 1) et temporelles
- Retrait des valeurs exceptionnelles (*outlierKD*)

Voyelle	i	ɚ	ɔ	o
n =	7002	3980	5749	6182
F1 ()	444 Hz	459 Hz	518 Hz	514 Hz
F2 ()	1911 Hz	1744 Hz	1472 Hz	1584 Hz



Hypothèses

D'après la littérature (Stuart-Smith *et al.* 2015) :

- /i/ (FLEECE) est stable
- /ʌ/ (FOOT/GOOSE) descend
- /ɔ/ (LOT) monte et devient plus antérieur
- /o/ (GOAT) monte



Hypothèses

D'après la littérature (Stuart-Smith *et al.* 2015) :

- /i/ (FLEECE) est stable
 - F1 et F2 stables
- /ʊ/ (FOOT/GOOSE) descend
 - F1 monte, F2 stable
- /ɔ/ (LOT) monte et devient plus antérieur
 - F1 descend, F2 monte
- /o/ (GOAT) monte
 - F1 descend, F2 stable



Analyse des données en diachronie/synchronie

- Variation en temps réel et apparent : analyse statistique en régression linéaire à effets mixtes
- Un modèle/voyelle avec F1/F2 en variable principale
- Lieu d'articulation à droite et à gauche, groupe d'âge et genre en facteurs fixes
- Mot et locuteur en facteurs aléatoires



Analyse des données en dynamique

- Variation dynamique intra et inter-locuteurs
- Méthodes utilisées par le passé : paradigme AXB, analyse discriminante, moyennes glissantes, GAMMs
- Nouvelle option : la méthode d'Alexander (2018), analyse *lmer* prenant en compte la dimension temps



II – AVANCEMENT DU PROJET : RESULTATS

Analyse en diachronie/synchronie

/i/ FLEECE			
F1	90M-90Y	*	↗
	90M-00M	**	↗
F2	90M-00M	**	↗
	00M-00Y	***	↗

/ʊ/ FOOT, GOOSE			
F1	90M-90Y	**	↗
	00M-00Y	***	↗
F2	90M-00M	**	↘

/ɔ/ LOT			
F1	90M-90Y	*	↗
	90M-00M	**	↗
F2	90M-00M	**	↗
	00M-00Y	***	↗

/o/ GOAT			
F1	90M-90Y	***	↗
	90M-00M	**	↗
	90Y-00Y	**	↘
F2	90M-00M	**	↘
	90Y-00Y	**	↘

II – AVANCEMENT DU PROJET : RESULTATS

Analyse en diachronie/synchronie

/i/ FLEECE				
F1	90M-90Y *	↗	X	
	90M-00M **	↗		
F2	90M-00M **	↗	X	
	00M-00Y ***	↗		

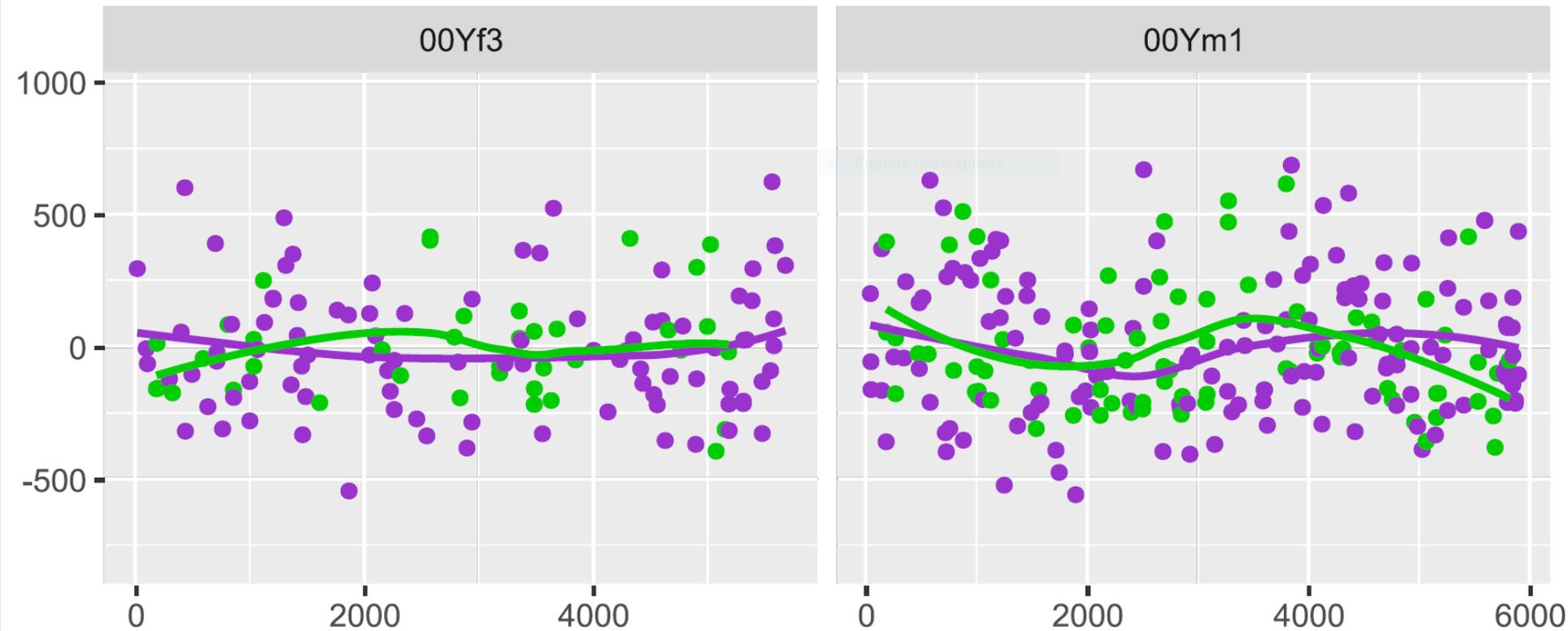
/ʊ/ FOOT, GOOSE				
F1	90M-90Y **	↗	✓	
	00M-00Y ***	↗		
F2	90M-00M **	↘	X	

/ɔ/ LOT				
F1	90M-90Y *	↗	X	
	90M-00M **	↗		
F2	90M-00M **	↗	✓	
	00M-00Y ***	↗		

/o/ GOAT				
F1	90M-90Y ***	↗	?	
	90M-00M **	↗		
	90Y-00Y **	↘		
F2	90M-00M **	↘	X	
	90Y-00Y **	↘		

II – AVANCEMENT DU PROJET : RESULTATS

Analyse de la variation dynamique (ici ex du F2 de /o/):



x 152

Analyse de la variation dynamique

- Aucune tendance visible par voyelle/formant, contrairement aux hypothèses.
- Mais différents comportements d'accommodation ?



Analyse de la variation dynamique

- Aucune tendance visible par voyelle/formant, contrairement aux hypothèses.
- Mais différents comportements d'accommodation ?
 - Absence de variation
 - Variation synchronisée
 - Variation asynchrone



Conclusions

- Y a-t-il un lien entre convergence à court terme et changement à long terme ?
- Y a-t-il des situations plus ou moins propices à la convergence ? Euh ... Si, si, le genre ! Enfin je sais pas, enfin peut-être.



Conclusions

- Y a-t-il un lien entre convergence à court terme et changement à long terme ? **Euh ...**
- Y a-t-il des situations plus ou moins propices à la convergence ?



Conclusions

- Y a-t-il un lien entre convergence à court terme et changement à long terme ? **Euh ...**
- Y a-t-il des situations plus ou moins propices à la convergence ? **Euh ...**



Conclusions

- Y a-t-il un lien entre convergence à court terme et changement à long terme ? **Euh ...**
- Y a-t-il des situations plus ou moins propices à la convergence ? **Euh ... Si, si, le genre ! Enfin je sais pas, enfin peut-être.**



III – DIFFICULTES DU PROJET

Imperfections des ressources et
limitations externes



III – DIFFICULTES RENCONTREES : SEGMENTATION

Conclusions techniques du pilote :

Les valeurs temporelles résultants de l'alignement automatique ne sont pas acceptable, en raison du palier de segmentation par 10ms.



Conclusions techniques du pilote :

Les valeurs temporelles résultants de l'alignement automatique ne sont pas acceptable, en raison du palier de segmentation par 10ms.

- Pour conserver l'étude de la **quantité vocalique**, il faudra trouver un autre moyen de segmenter /i ʌ/ (voyelles sujettes à la *Scottish Vowel Length Rule*, soit l'alternance brèves/longues selon le contexte morpho-phonétique).



Conclusions techniques du pilote :

Les valeurs temporelles résultants de l'alignement automatique ne sont pas acceptable, en raison du palier de segmentation par 10ms.

- Pour améliorer l'étude de la **qualité vocalique**, il faudra sûrement affiner la fenêtre d'extraction des valeurs spectrales (0.25 et 0.75 au lieu de 0 et 1 ?)



III – DIFFICULTES RENCONTREES : TEMPUS FUGIT, ET SOLUM TRES ANNOS HABES

Conclusions du pilote :

L'annotation manuelle de certains facteurs : un procédé très chronophage inautomatisable

- Facteurs prosodiques
- Facteurs conversationnels



IV – POURSUITE DU PROJET

Surmonter les difficultés,
dépasser le premier pilote,
triompher de la thèse



Démarrer les prochaines étapes prévues :

- Ajouter les autres segments vocaliques
- De l'intégralité du corpus (autres types d'interactions, 8 groupes d'âge restants)
- Annotés selon les facteurs sociaux et situationnels
- Et selon davantage de facteurs linguistiques



Grâce à **ISCAN** (*Integrated Speech Corpus Analysis*, McAuliffe 2019), formation à l'automne 2019 : possibilité d'étudier à grande échelle la réalisation de variables

- Consonantiques (délai de voisement, /r/ et /l/, rétraction des sibilantes ...)
- Suprasegmentales (débit de parole, fréquence fondamentale ...)



IV – POURSUITE DU PROJET

Mais aussi ne pas oublier de dire merci !

