



UNIVERSITÉ DE GENÈVE 

Ce que les corpus PFC nous apprennent sur la phonologie prosodique du français

Mathieu Avanzi, Université de Neuchâtel
Sandra Schwab, Université de Genève


Avanzi PFC - Paris 21/2011, 26/12/2012, décembre 2012

UNIVERSITÉ DE GENÈVE 

Objectifs de la communication

- Présenter une partie des résultats d'une recherche en cours portant sur la **variation prosodique régionale en français de Suisse romande**
- **4 heures** du corpus **PFC**, échantillonnées pour tenir compte de la **variété** (région); de l'**âge**; du **sexe**; du **style de parole** (lecture vs conversation)


Avanzi et al. [2011, 2012]; Avanzi, Schwab et al. [2012]; Obin et al. [2012]; Schwab et al. [à par.]

UNIVERSITÉ DE GENÈVE 

Objectifs de la communication


- Quelle **validité empirique** pour les **régularités** et les **règles phonologiques** décrites dans la littérature sur la prosodie du français?
- Zoom sur les phénomènes de **non-accentuation** lexicale phonologiquement « contraints »: **Align-XP** et ***CLASH**

Avanzi, [à par.], Avanzi & Schwab [en prép.]


UNIVERSITÉ DE GENÈVE 

Plan de la communication

1. L'accentuation du français
2. Corpus
3. Résultats
4. Conclusion

UNIVERSITÉ DE GENÈVE 


1. L'accentuation du français

UNIVERSITÉ DE GENÈVE 

L'accentuation du français
l'accent primaire (ou final)

- En français, domaine de l'**accent primaire**
 - ≠ mot lexical mais = **groupe clitique**
 - Pas de fonction distinctive (≠ langues romanes)
- L'accent tombe sur le **bord droit** de groupes clittiques
 - Syllabe masculine (dont le noyau n'est pas un schwa)
 - Accent primaire = **frontière** → **accent final**
 - **Synchrétisme** entre accentuation et intonation
- Conséquences
 - Français langue « **sans accent** »
 - Français langue « **à frontière** »

Garde (1969); Naimo (2004); Rossi (1975); Vaisse (1990); Lachet & Besengetre (1999)


UNIVERSITÉ DE GENÈVE 

L'accentuation du français

le groupe accentuel (GA)

- Un groupe clitique = un groupe accentuel « **potentiel** »
- Dans la parole, tous les groupes clitiques ne sont pas assortis d'un **accent final**
 - ⇒ Seuls les groupes clitiques assortis d'un accent génèrent un **groupe accentuel**
 - ⇒ Pas d'isomorphisme strict entre groupe clitique et groupe accentuel
- Groupe accentuel ≙
 - Mot prosodique
 - Mot phonologique
 - Syntagme accentuel (AP)
 - Syntagme phonologique (φ)
 - Groupe intonatif

Selkirk [1978], Martin [1987], Merens [1989], Delais-Roussarie [1996], Jun & Fougeron [2002]


UNIVERSITÉ DE GENÈVE 

L'accentuation du français

phénomènes de non-accentuation – généralités

- Plusieurs raisons peuvent expliquer la **non-accentuation** de la syllabe finale des groupes clitiques
 - Contraintes de **débit** : plus on parle vite, moins on réalise d'accents en fin de groupe clitique
 - Contraintes **sémantiques** : désaccentuer un élément pour signaler son inclusion dans un syntagme de rang supérieur
- Certains phénomènes de non-accentuation s'expliquent autrement, car ils sont **codés dans la phonologie**

Fougeron & Jun [1998], Post [2011]


UNIVERSITÉ DE GENÈVE 

L'accentuation du français

phénomènes de non-accentuation – ALIGN-XP

- **ALIGN-XP**
Les adjectifs (et autres modifieurs) ne génèrent pas d'accent final quand ils sont antéposés au nom qu'ils modifient
 - les jolis pantalons
 - (les jolis)_{GC} (pantalons)_{GC} >> (les jolis pantalons)_{GA}

Verhuyten [1982], Selkirk [1984], Nespor & Vogel [1986], Delais-Roussarie [1996], Post [2000]


UNIVERSITÉ DE GENÈVE 

L'accentuation du français

phénomènes de non-accentuation – *CLASH

- ***CLASH**
Deux syllabes adjacentes ne peuvent pas être toutes deux assorties d'un accent primaire.
 - on connaît Paul
 - (on connaît)_{GC} (Paul)_{GC} >> (on connaît Paul)_{GA}

Verhuyten [1982], Selkirk [1984], Dell [1984], Nespor & Vogel [1986], Delais-Roussarie [1996], Post [2000]

UNIVERSITÉ DE GENÈVE 

L'accentuation du français


synthèse et problématique

- Pas d'études sur **corpus échantillonnés** qui permettraient de juger de la validité empirique de tels faits
- Résultats obtenus à l'issue d'examens de données lues, mais faisant office d'**acquis** en phonologie
 - Quid de la **variation régionale**?
 - Quid de la **parole « ordinaire »**?

Lafrenay-Dugnat [2003], Arai [2002], Cole et al. [2012]

UNIVERSITÉ DE GENÈVE 

2. Matériel

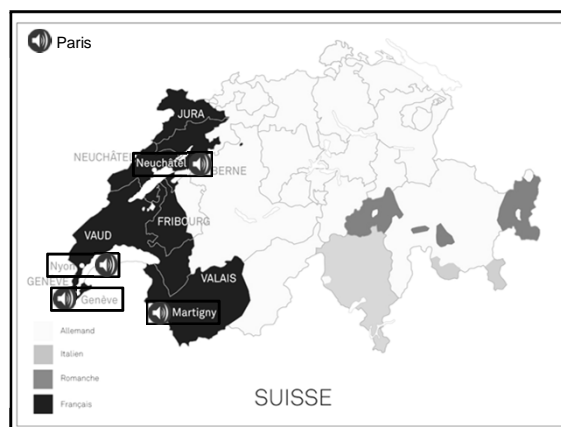
UNIVERSITÉ DE GENÈVE 


Corpus

- 5 points d'enquête
- 8 locuteurs par point d'enquête
 - Sexe et âge contrôlés
- 2 styles de parole par locuteur
 - Lecture du texte PFC (180 ms)
 - Extrait monologique d'une conversation (180 ms)

Andreasen & Ljch, [2009], Avanzi, Schwab et al. [2012], Racine & Andreasen [2012]

13




UNIVERSITÉ DE GENÈVE 

Corpus

- 5 points d'enquête
- 8 locuteurs par point d'enquête
 - Sexe et âge contrôlés
- 2 styles de parole par locuteur
 - Lecture du texte PFC (180 ms)
 - Extrait monologique d'une conversation (180 ms)


15

UNIVERSITÉ DE GENÈVE 

Corpus

- 4 hommes et 4 femmes
 - Paris (PA) entre 24 et 86 (moyenne 50.3)
 - Genève (GE) entre 21 et 59 (moyenne 41.3)
 - Nyon (NY) entre 20 et 82 (moyenne 43.6)
 - Martigny (MA) entre 22 et 78 (moyenne 48.8)
 - Neuchâtel (NE) entre 25 et 78 (moyenne 52.5)
- Âge = variable contrôlée
 - Région (F (4, 30) = 0,26, n.s.)
 - Sexe (F (1, 30) = 0,05, n.s.)
 - Interaction Région*Sexe (F (4, 30) = 0,11, n.s.)

16


UNIVERSITÉ DE GENÈVE 

Corpus

- 5 points d'enquête
- 8 locuteurs par point d'enquête
 - Sexe et âge contrôlés
- 2 styles de parole par locuteur
 - Lecture du texte PFC (~180 ms)
 - Extrait monologique d'une conversation (~180 ms)

Davies et al. [2012, 2010]

17

UNIVERSITÉ DE GENÈVE 


Corpus

- Au total

Nb_Syll	Nb_Prom	Nb_H	Nb_NProm	Nb_GA	%Prom_F	%Prom_NF
49'398	17'695	2'116	29'587	14'606	30,867	6,529

– durée ~ 4h00


18

UNIVERSITÉ DE GENÈVE 

Transcription et codage

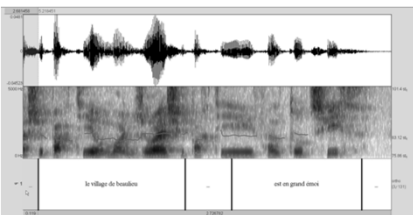
1. Transcription orthographique
2. Alignement phonèmes/syllabes/mots
3. Codage des proéminences
4. Segmentation en groupes accentuels
5. Identification des sites d'application des règles phonologiques ALIGN-XP et *CLASH

Avanzi, Bordin & Olin (2011); Avanzi (2013)


UNIVERSITÉ DE GENÈVE 

1. Transcription orthographique

[Boersma & Weenink 2012]

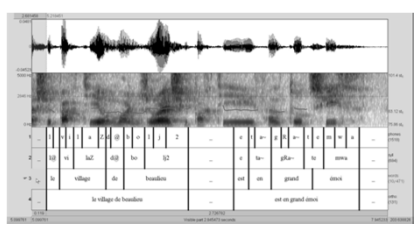


le village de bondeux est en grand bois


UNIVERSITÉ DE GENÈVE 

2. Alignement phones/syll./mots

[Goldman 2011]

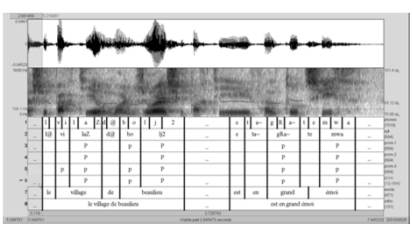


le village de bondeux est en grand bois

UNIVERSITÉ DE GENÈVE 


3. Codage des proéminences

[Avanzi et al. 2007, 2010, à par.]

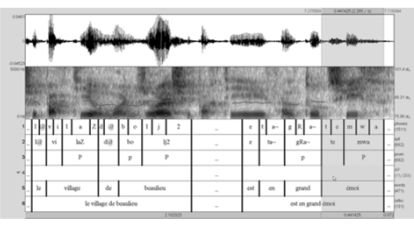


le village de bondeux est en grand bois


K = 0.65/0.91

UNIVERSITÉ DE GENÈVE 

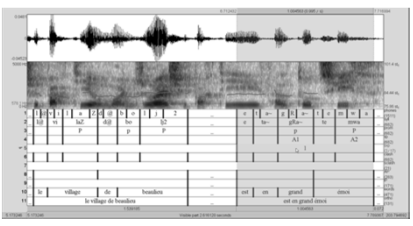
4. Segmentation en GA



le village de bondeux est en grand bois

UNIVERSITÉ DE GENÈVE 

5. Identification des sites ALIGN-XP et *CLASH



le village de bondeux est en grand bois

UNIVERSITÉ DE GENÈVE unine

Statistiques

prédicteurs et interactions

- **Analyse statistique**
 - Modèle linéaire généralisé à mesures répétées
- **Prédicteurs**
 - Variété (GE, MA, NE, NY, PA)
 - Style de parole (Conversation vs Lecture)
 - Vitesse d'articulation (= durée syllabique)
- **Interactions prises en compte**
 - Variété*Style de parole
 - Variété*Vitesse d'articulation

25

UNIVERSITÉ DE GENÈVE unine

3. Résultats

UNIVERSITÉ DE GENÈVE unine

L'accentuation du français

phénomènes de non-accentuation – ALIGN-XP

- **ALIGN-XP**
Les adjectifs (et autres modificateurs) ne génèrent pas d'accent final quand ils sont antéposés au nom qu'ils modifient

N = 443

- est en grand émoi
- (est en grand)_{GC} (émoi)_{GC} >> (est en grand émoi)_{GA}

Verhuyten [1982], Selkirk [1984], Nespor & Vogel [1986], Delais-Roussarie [1996], Post [2000]

27

UNIVERSITÉ DE GENÈVE unine

Align-XP

variété

Wald $\chi^2(4) = 4.611, n.s.$

Figure 1: Pourcentage de respect de la règle ALIGN-XP (estimé par le modèle) selon la variété. Les barres d'erreur correspondent à l'erreur standard de la moyenne.

28

UNIVERSITÉ DE GENÈVE unine

Align-XP

style de parole

Wald $\chi^2(4) = 22.995, p < 0.001$

Figure 2: Pourcentage de respect de la règle ALIGN-XP (estimé par le modèle) selon le style de parole. Les barres d'erreur correspondent à l'erreur standard de la moyenne.

29

UNIVERSITÉ DE GENÈVE unine

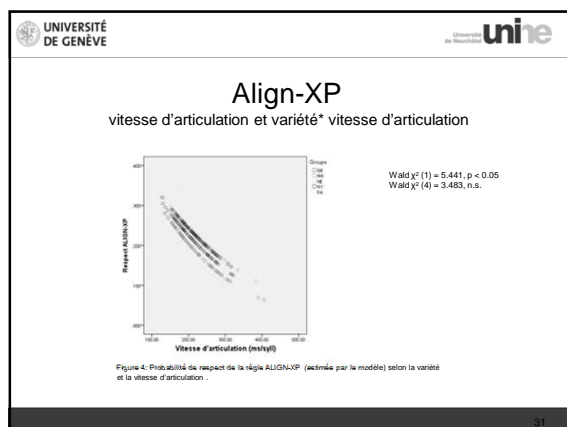
Align-XP

variété*style de parole

Wald $\chi^2(4) = 5.057, n.s.$

Figure 3: Pourcentage de respect de la règle ALIGN-XP (estimé par le modèle) selon la variété et le style de parole. Les barres d'erreur correspondent à l'erreur standard de la moyenne.

30



UNIVERSITÉ DE GENÈVE **unine**

Respect de la règle Align-XP

résumé

- Pas d'effet de **région**
- Effet du style de **parole**
 - Règle mieux respectée en conversation qu'en lecture
- Effet de **vitesse d'articulation**

32

UNIVERSITÉ DE GENÈVE **unine**

L'accentuation du français

phénomènes de non-accentuation – *CLASH

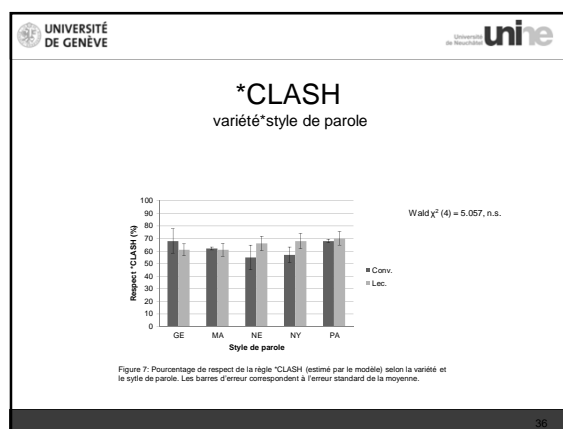
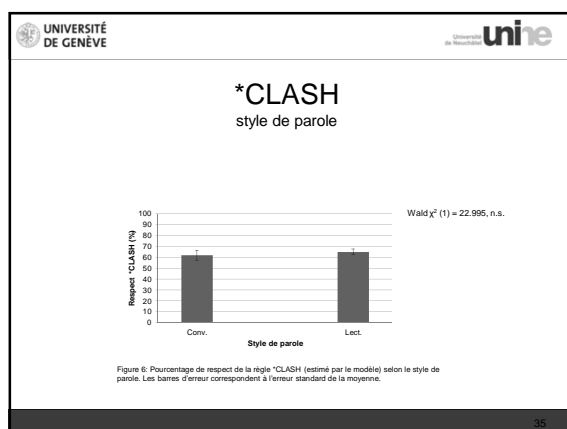
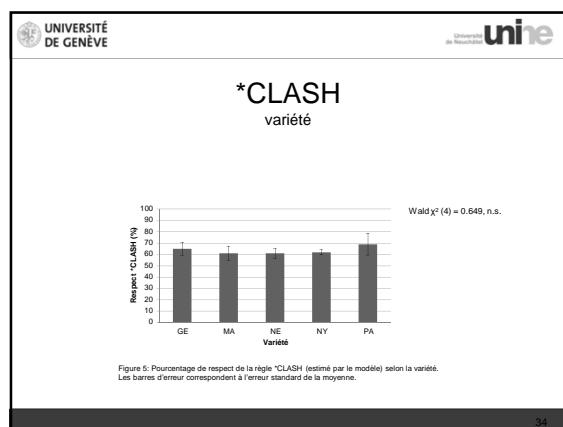
- ***CLASH**
Deux syllabes adjacentes ne peuvent pas être toutes deux assorties d'un accent primaire.

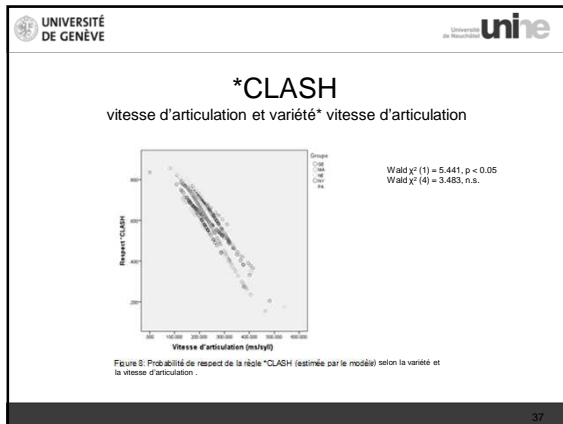
– on connaît Paul
– (on connaît)_{GC} (Paul)_{GC} >> (on connaît Paul)_{GA}


N = 576

Verkuylen [1952]; Selkirk [1984]; Dell [1984]; Nespor & Vogel [1986]; Delais-Roussarie [1996]; Post [2000]

33






UNIVERSITÉ DE GENÈVE 

*CLASH

résumé

- Pas d'effet de **région**
- Pas d'effet du **style de parole**
- Effet de **vitesse d'articulation**

38

UNIVERSITÉ DE GENÈVE 


Discussion

ALIGN-XP

- **Très peu respectée** (25% de respect en moyenne)
- Pas sensible à la **variété**
- Mais sensible au **style de parole**
 - Mieux respectée en conversation qu'en lecture
 - Pourquoi?
- Et à la **vitesse d'articulation**
 - Plus on parle lentement, plus on a de chances de violer la règle
- Des mesures complémentaires montrent que
 - La concomitance d'ALIGN-XP avec *CLASH **augmente** la probabilité de respect
 - Le **sémantisme** de l'adjectif joue un rôle : la probabilité de **viol augmente** pour les adj. marqués référentiellement (*gros, grands*)
 - Quid de la **fréquence** des adjectifs?

Thalier (2012); Avanzi & Schwab (en prép.)

39


UNIVERSITÉ DE GENÈVE 

Discussion

*CLASH


- Mieux respectée que ALIGN-XP (plus de 60% de respect en moyenne)
- Ni sensible à la **variété**, ni au **style de parole**
- Mais sensible à la **vitesse d'articulation**
 - Plus on parle lentement, plus on a de chances de violer la règle

40

UNIVERSITÉ DE GENÈVE 

4. Conclusion


41

UNIVERSITÉ DE GENÈVE 

Résumé

- **Règles phonologiques** de non-accentuation lexicale en français
 - Approche sur **corpus** (4 heures de données échantillonnées pour la variation régionale, en tenant compte du style de parole, de l'âge et du sexe des locuteurs)
 - Approche **assistée informatiquement** (alignements, codages et calculs facilités par des scripts, ce qui permet de traiter de « grandes » masses de données)
- Principaux résultats
 - Règles de non-accentuation sont plus ou moins bien respectées
 - Influence nulle de la **variation régionale**
 - Influence mitigée du **style de parole**
 - Influence importante de la **vitesse d'articulation**


42

UNIVERSITÉ DE GENÈVE 

Discussion

- Questions en suspens:
 - À partir de quel **pourcentage de réussite** peut-on ériger **en règle phonologique** une contrainte?
 - **Sensibilité** à la vitesse d'articulation : sensibilité du phénomène → quel statut phonologique ?

43

UNIVERSITÉ DE GENÈVE 

Ce que les corpus PFC nous apprennent sur la phonologie prosodique du français

Mathieu Avanzi, Université de Neuchâtel
Sandra Schwab, Université de Genève

Corpus PFC – Paris, FUSP, 16/17 décembre 2012