

Variations dans la formation du syntagme accentuel en français

Analyse de 5 points d'enquête PFC

Guri Bordal, Université Paris Ouest/Universitet i Oslo

Mathieu Avanzi, Université de Neuchâtel

Nicolas Obin, IRCAM/CNRS Paris

Alice Bardiaux, FNRS/Université catholique de Louvain

Introduction

- **Objectif** : Examiner la réalisation du syntagme accentuel dans les variétés de français en contact avec les langues à prosodie lexicale

- Cinq variétés étudiées

- Deux variétés **[-contact]**

- Français parisien [FR-76]
- Français neuchâtelois [SW-NE]



- Trois variétés **[+contact]**

- Français centrafricain [AF-CFA]
- Français sénégalais [AF-SN]
- Français neuchâtelois [SW-GER]
parlé par des
germanophones



La typologie prosodique

	+ ton lexical	- ton lexical
+ accent lexical	Norvégien, japonais	Wolof, allemand
- accent lexical	Sango, bambara, chinois	Français, coréen

L'hypothèse de marque typologique

- La majorité des langues ont une prosodie lexicale;
- Les langues comme le français [-accent lexical, -ton lexical] sont rares;
- La prosodie du français est plus marquée que la prosodie des langues à prosodie lexicale

⇒ **Prédiction** : un système typologiquement marqué disparaît en contact avec des systèmes moins marqués

Les variétés [+contact]

- Le français centrafricain
 - Locuteurs de L1 le sango [+ton lexical]
- Le français sénégalais
 - Locuteurs de L1 le wolof [+accent lexical]
- Le français neuchâtelois germanophone
 - Locuteurs de L1 suisse allemand [+accent lexical]

Le syntagme accentuel

- La plus petite unité prosodique du français
- Domaine de l'accent lexical
- Correspond *a minima* au groupe clitique (un mot lexical et les mots fonctionnels qui en dépendent)
- Son poids métrique dépend de contraintes diverses (rythmiques, syntaxiques, sémantiques)

elle est venue me voir l'autre jour

⇒ (elle est venue) (me voir) (l'autre jour)

⇒ (elle est venue me voir) (l'autre jour)

⇒ (elle est venue me voir l'autre jour)

Le syntagme accentuel

- Deux contraintes sont considérées particulièrement puissantes dans sa formation
 - Rythmique : **NoClash**: (la voiture) (bleue) => (la voiture bleue)
 - Syntaxique : **Align-XP**: (les gentils) (enfants) => (les gentils enfants)
- Le SA est également le domaine des phénomènes de sandhi
 - **Liaison**: [lezãfã] vs *[leãfã]
 - **Enchaînement**: [mɛ.mã.ta.me] vs *[mɛm.ã.ta.me]

Hypothèses

- Les locuteurs des variétés [-contact] vont réaliser les syntagmes accentuels selon les prédictions du modèle
- Les locuteurs des variétés [+contact] vont segmenter la parole en unités plus petites, voire en unités lexicales
 - Bordal [2011]; Boutin & Gueye [à par.]

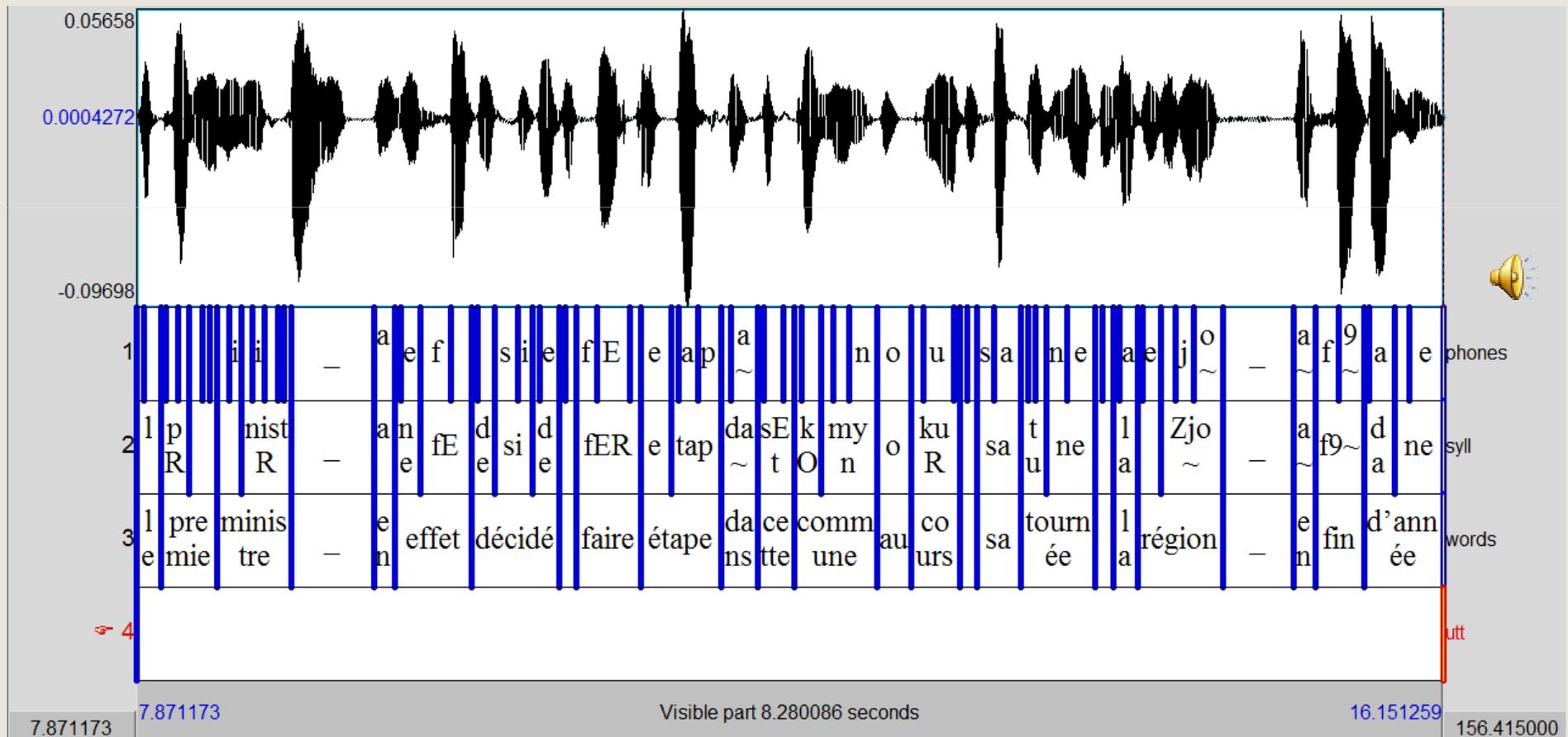
Données

- Le texte PFC lu
 - 22 phrases, 398 mots, +/- 130 sec.
- 4 locuteurs / variétés
 - 2 femmes et 2 hommes
 - Différentes tranches d'âge
- **Total**
 - 20 locuteurs
 - 43 minutes environ

Transcription/segmentation

1. Transcription orthographique sous **Praat**
2. Segmentation semi-automatique en mots graphiques, syllabes et phonèmes avec **EasyAlign**; vérification manuelle des alignements
3. Identification manuelle des **groupes clitiques** sur des critères **purement morphosyntaxiques**

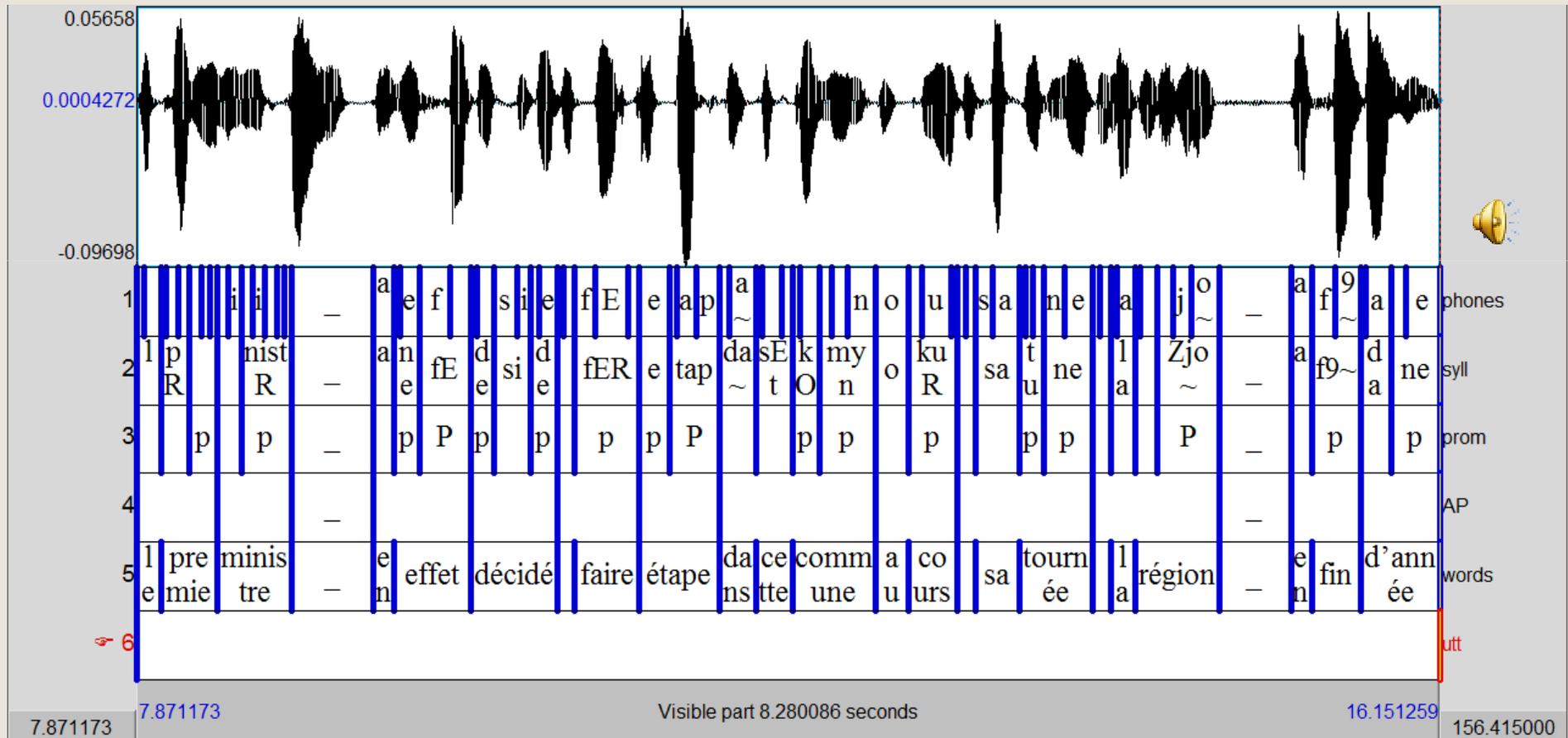
Segmentation



Méthode d'annotation prosodique

1. Détection des **syllabes proéminentes** sur la base de critères **perceptifs** par **deux** annotateurs experts
2. Création d'une tire **comparant** les annotations des experts
3. Un **troisième expert** intervient dans les cas de désaccord
4. Création d'une **tire de référence**

Annotation prosodique



Identification des sandhis

- Deux types de sandhis:
 - Liaisons
 - Consonne prononcée, pas de pause à droite
 - 🔊 [gRã.te.mwa]
 - 🔊 [gRã.e.mwa]
 - Enchaînement
 - Pas de pause/coup de glotte à droite de la consonne
 - 🔊 [fɛ.Rɛ.tap]
 - 🔊 [fɛR.e.tap]

Analyses

- Prise en compte de plusieurs types d'indices:
 1. Distribution des **proéminences**
 2. Respect de contraintes de **restructuration du SA** (NO_CLASH, ALIGN-XP)
 3. Domaine des **processus segmentaux (sandhis)**
- ⇒ Cf. Avanzi, Bordal & Obin (ICPHS, 2011), pour les origines de la méthodologie
- 4. Paramètres **rythmiques**
 5. Distribution des **silences**
- ⇒ Classification agglomérative non supervisée utilisée pour classifier les groupes d'après leurs caractéristiques moyennes

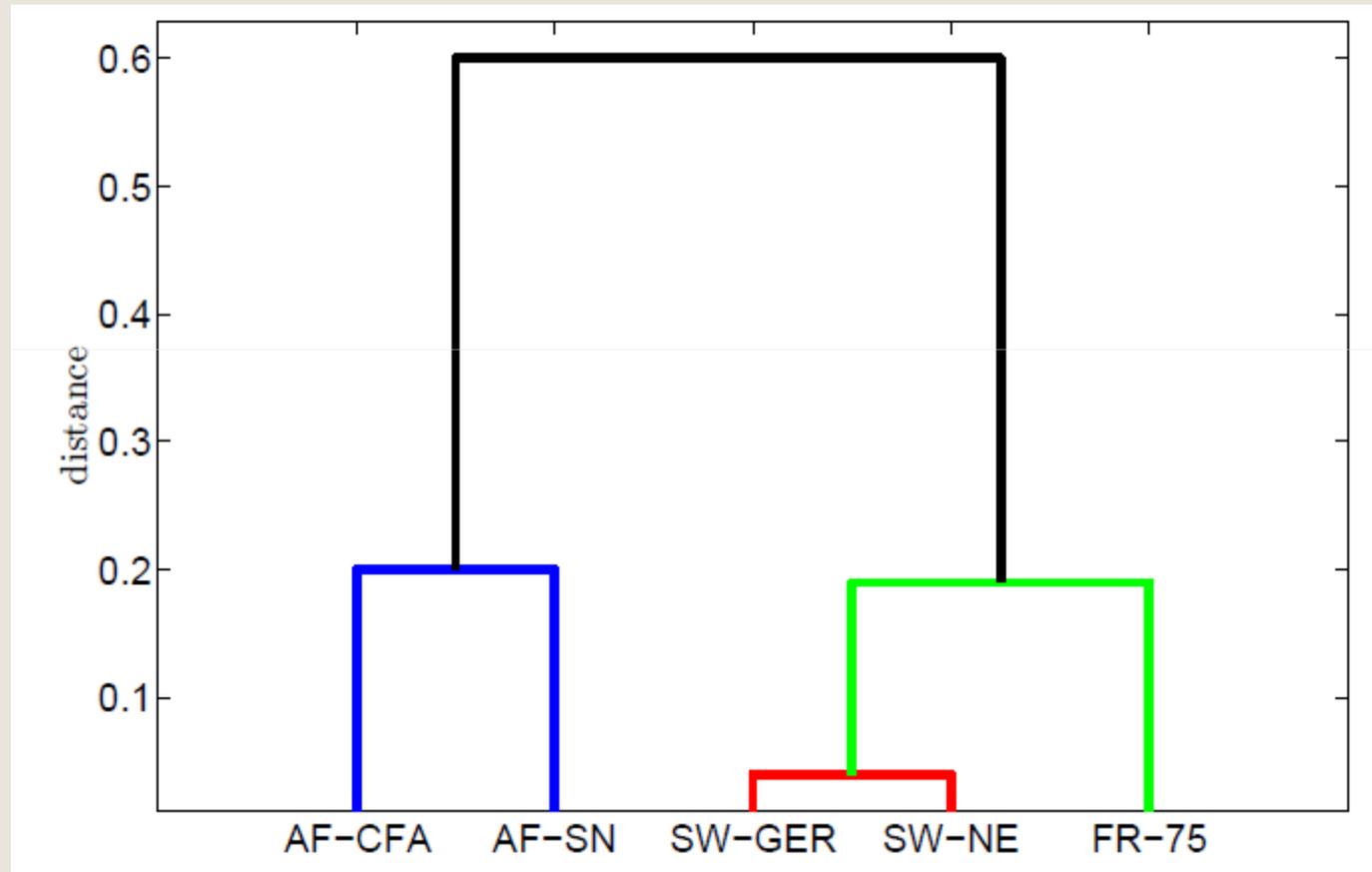
Vue d'ensemble

	FR-75		SW-NE		SW-GER		AF-SN		AF-CFA	
1. Prominences										
AP weight	3.4	(0.7)	3.1	(0.5)	3.1	(0.4)	3.1	(0.5)	2.8	(0.4)
Prominence ratio	35.0	(5.1)	41.8	(5.7)	46.2	(9.1)	43.3	(7.8)	46.5	(9.5)
2. AP restructuring										
Align-XP	43	(19)	1.62	(0.26)	71	(11)	69	(18)	57	7
Clash	18	(17)	36	(7)	82	(13)	57	(20)	89	(17)
3. Sandhis										
Sandhi_intern	100	(0)	98	(2)	74	(8)	83	(15)	66	(10)
Sandhi_extern	81	(10)	56	(7)	38	(6)	27	(13)	4	(8)
4. Rhythm										
Articulation rate	6.1	(0.5)	5.3	(0.5)	4.6	(0.6)	5.2	(0.7)	4.5	(0.8)
ΔC	4.1	(0.7)	5.3	(0.9)	5.4	(1.3)	6.3	(1.4)	6.2	(1.8)
%V	48.4	(5.8)	48.8	(4.8)	41.6	(6.6)	51.9	(6.9)	41.6	(7.9)
5. Silences										
Silence number	3.0	(2.1)	2.9	(2.1)	2.8	(1)	3.7	(2.9)	4.2	3.6
Silence duration (ms)	506	(143)	546	(233)	905	(453)	830	(300)	789	(210)

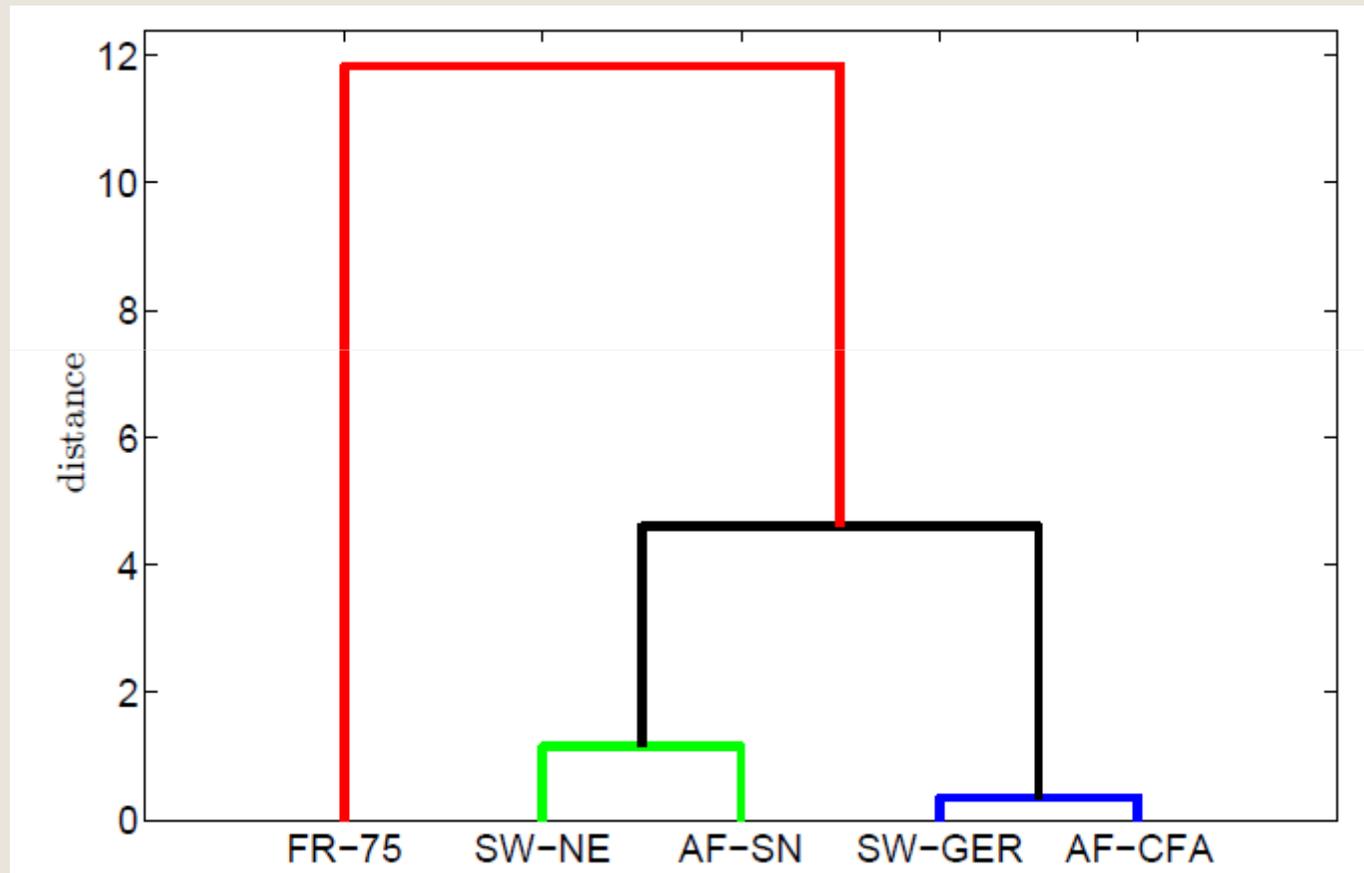
1- Distribution des proéminences

- Poids métrique de l'AP
- Ratio prom./syll.

Poids métrique de l'AP



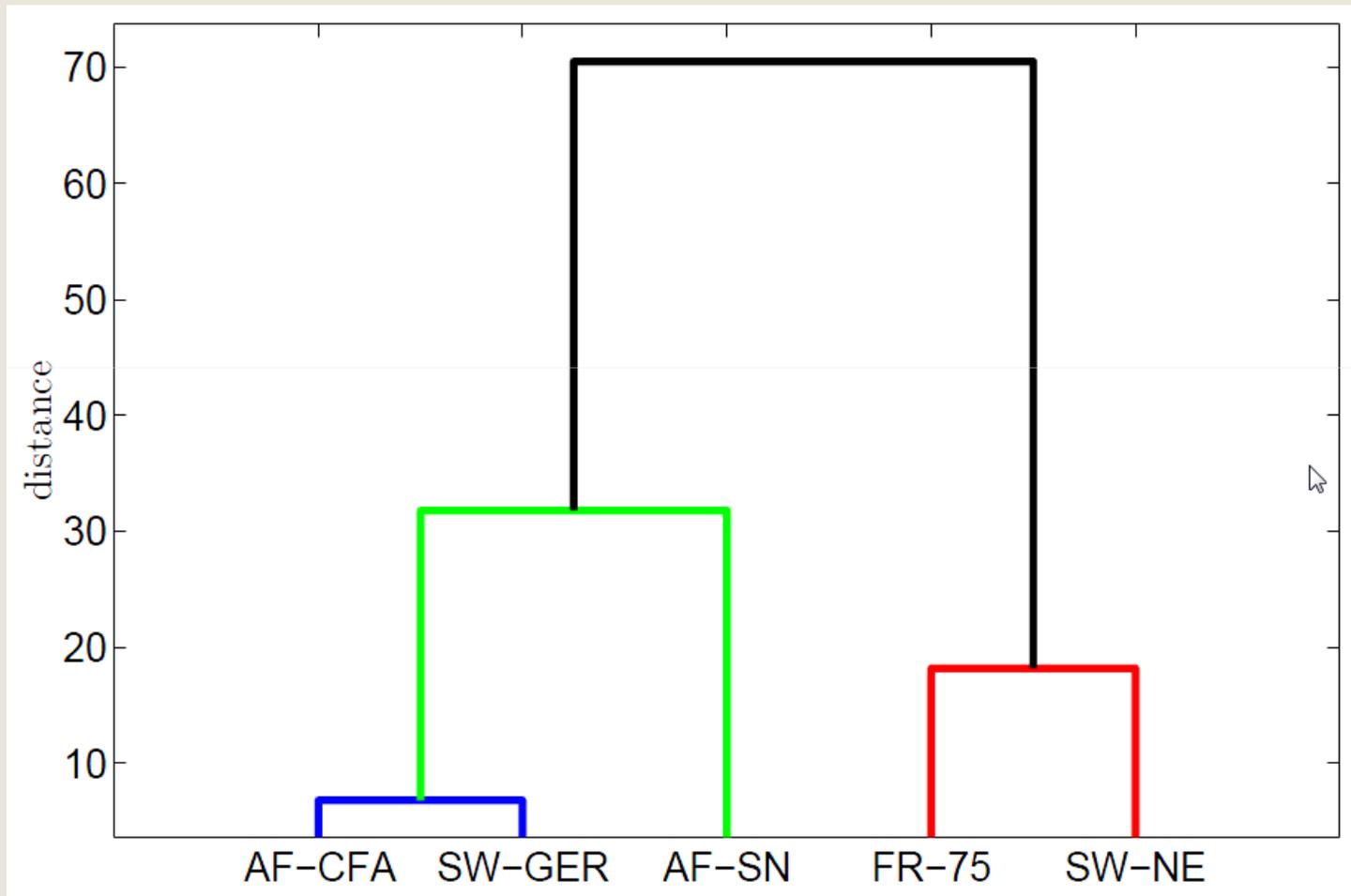
Ratio prom/syll.



2- Respect des règles de restructuration de l'AP

- ALIGN-XP
 - Aucune différence significative
- NO_CLASH

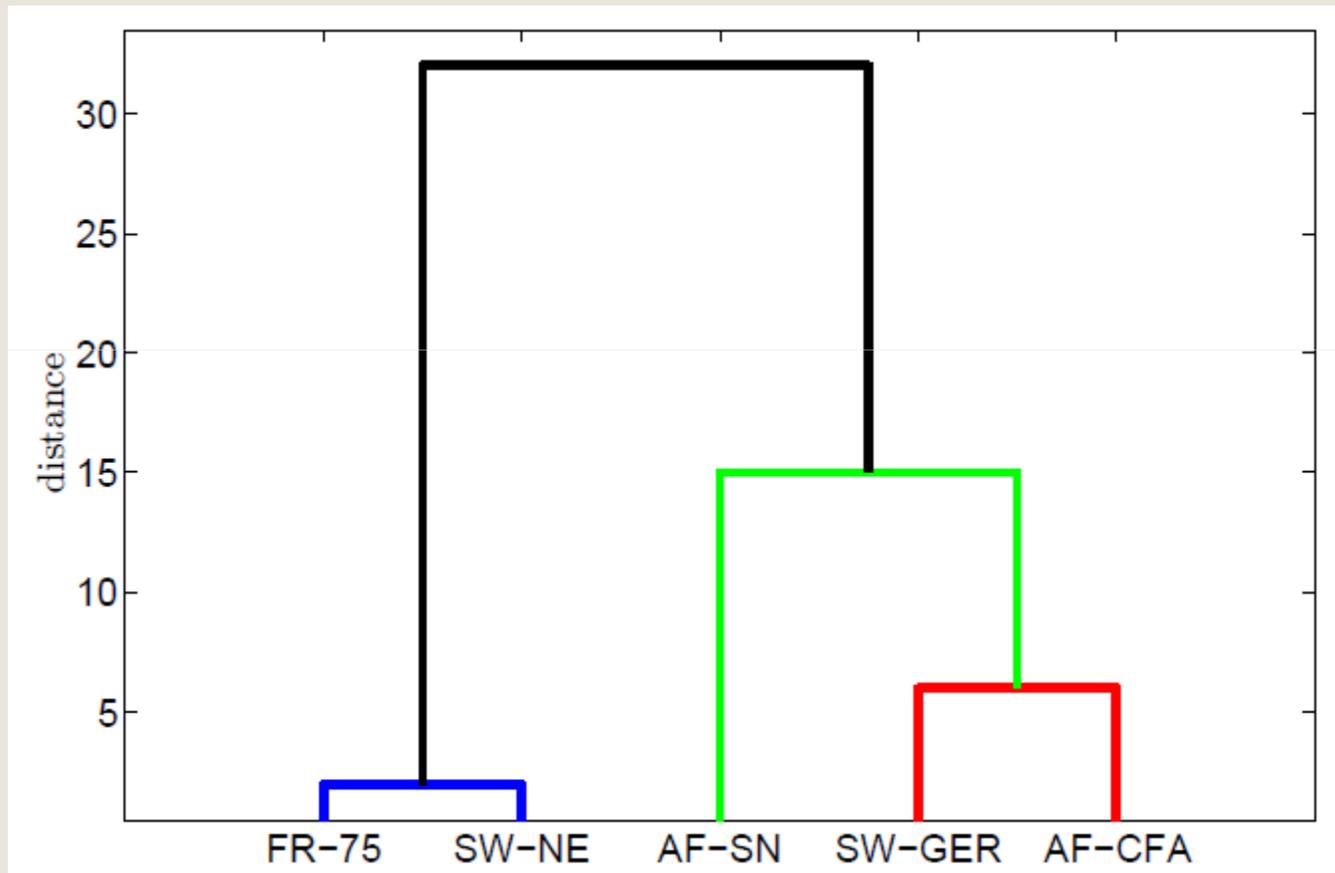
NO_CLASH



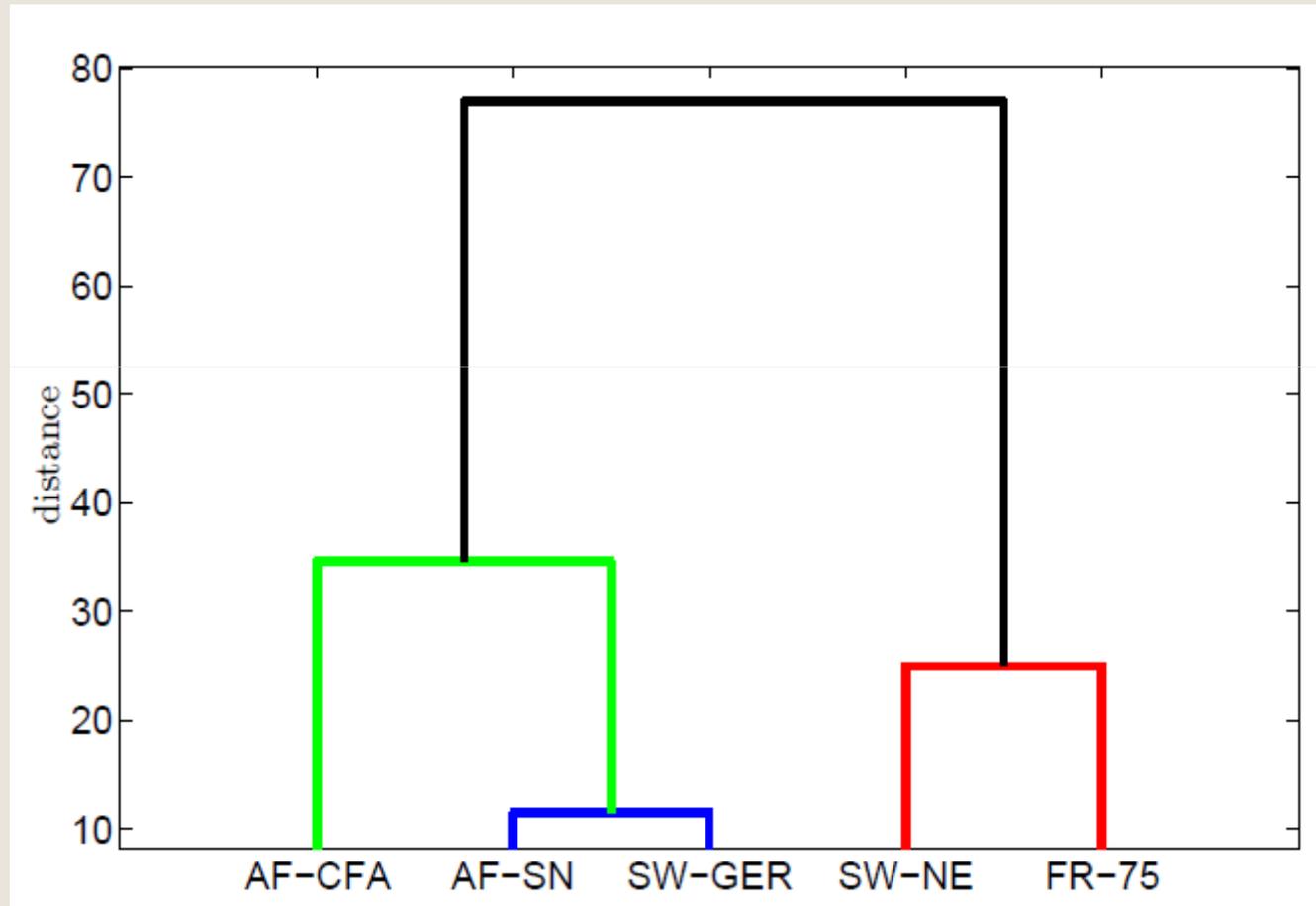
3- Distribution des sandhis

- Sandhis internes
- Sandhis externes

Sandhis internes



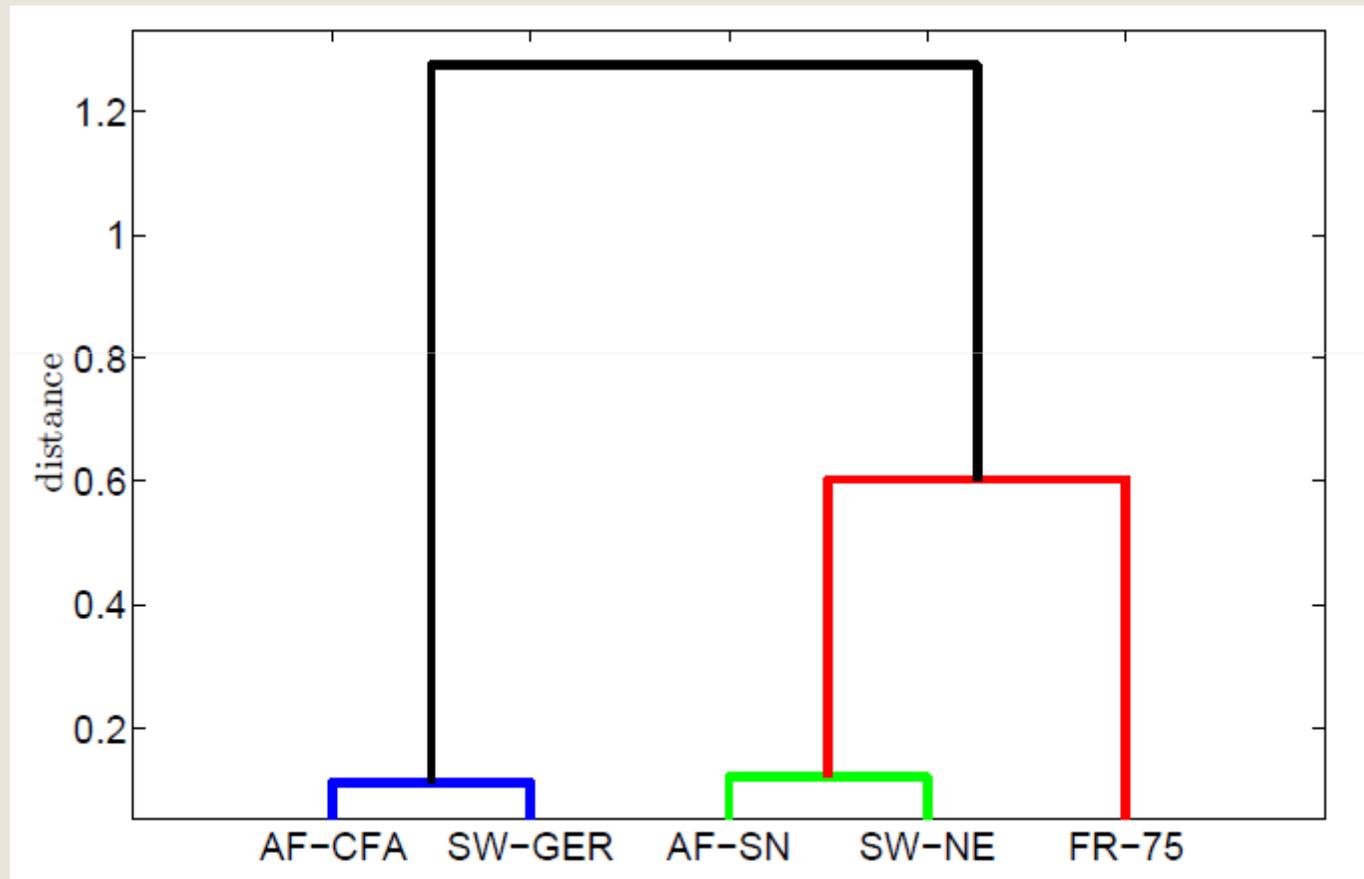
Sandhis externes



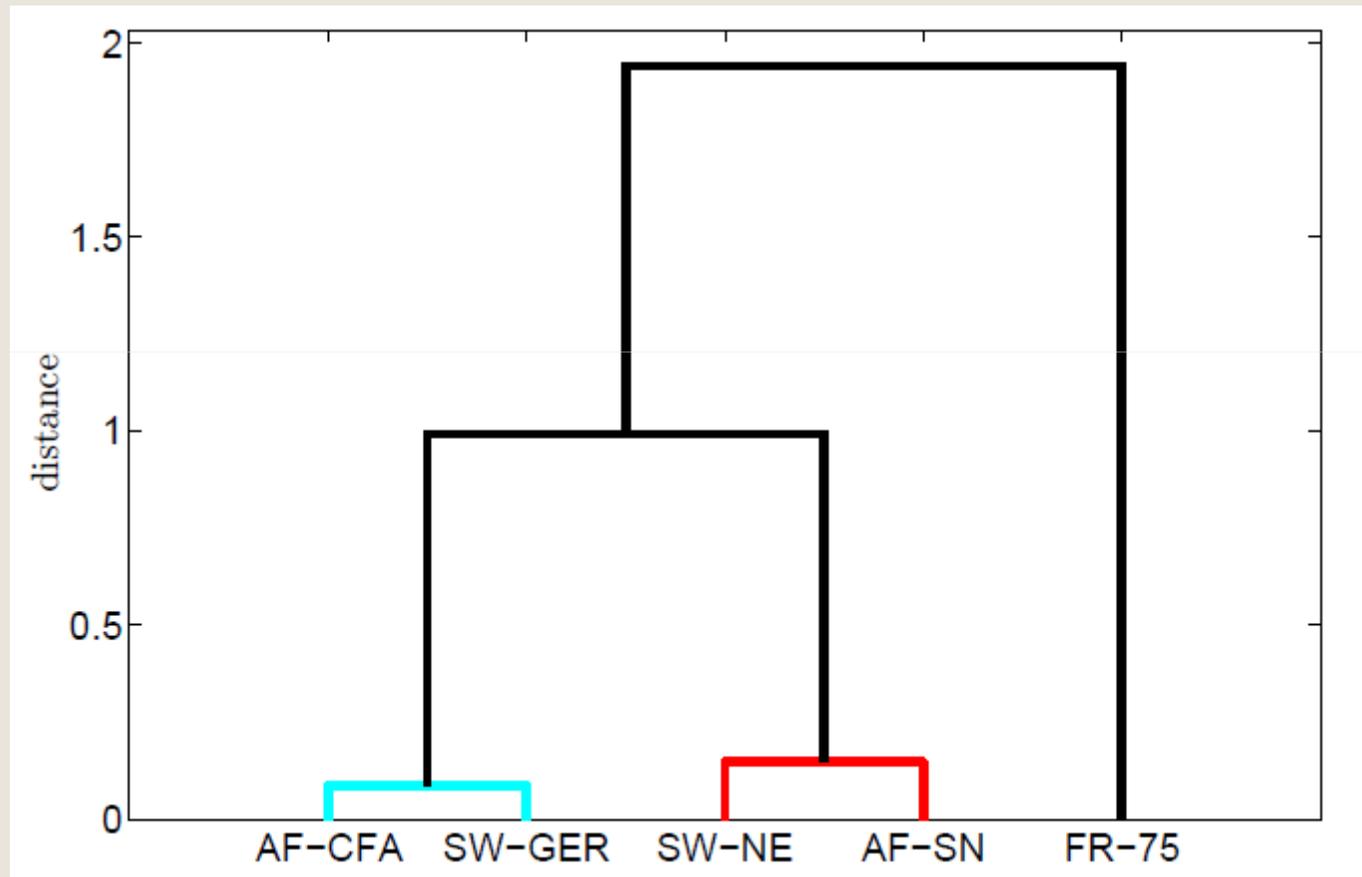
4- Paramètres rythmiques

- Taux d'articulation
- ΔC
- %V
 - Ne donne pas de résultats cohérents

Taux d'articulation



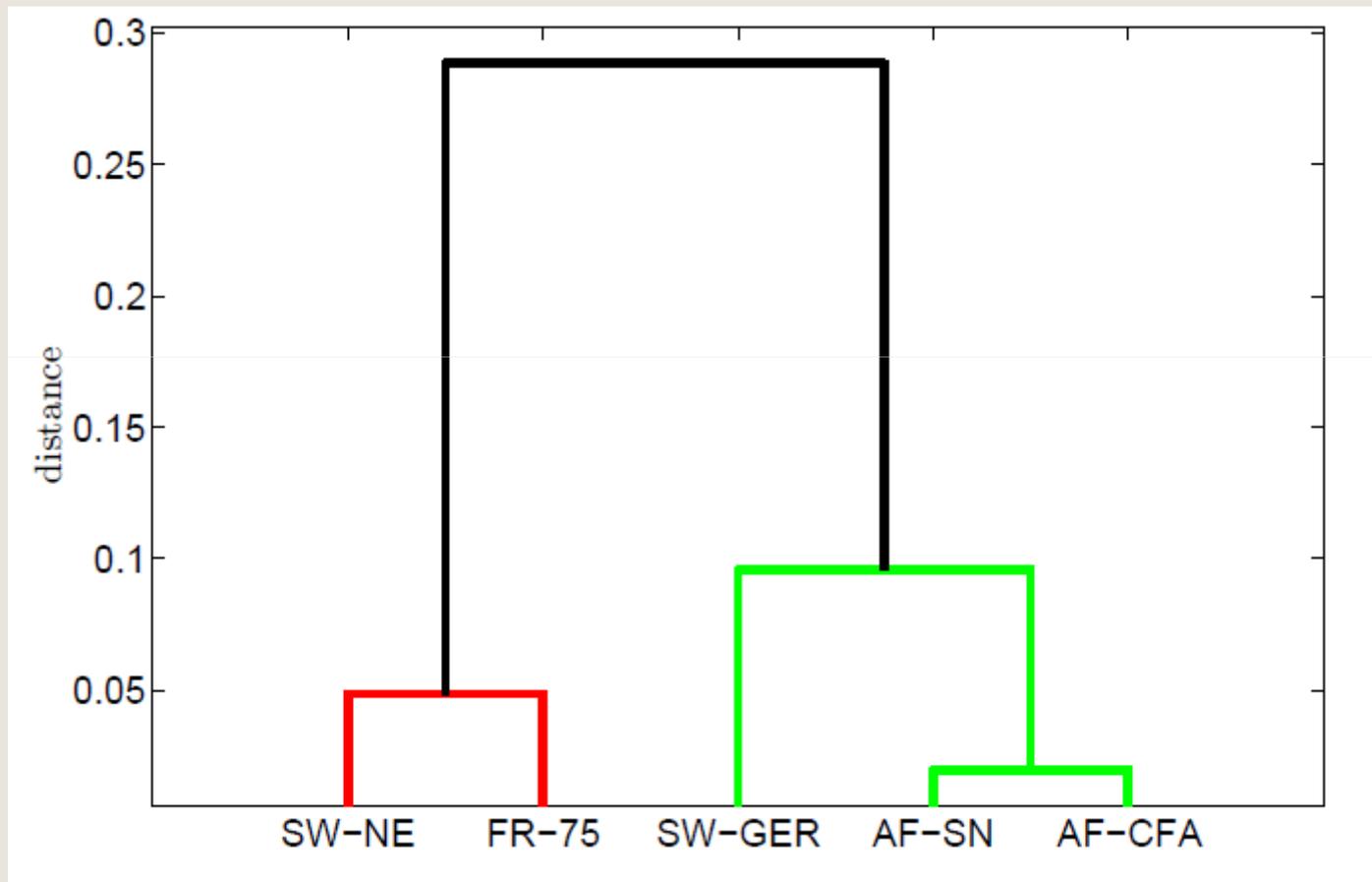
ΔC



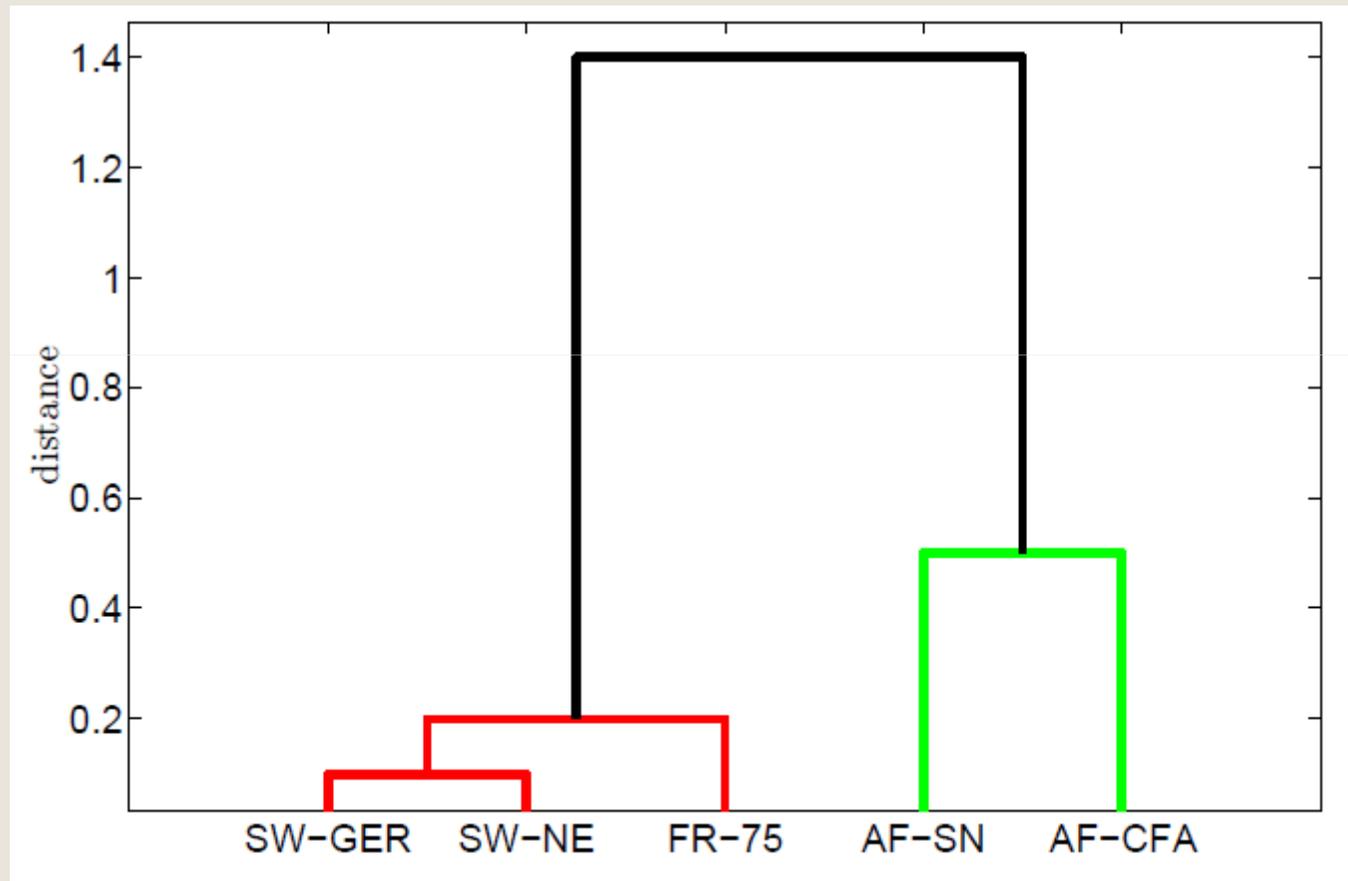
5- Distribution des silences

- Durée des silences
- Nombre des silences

Nombre de silences



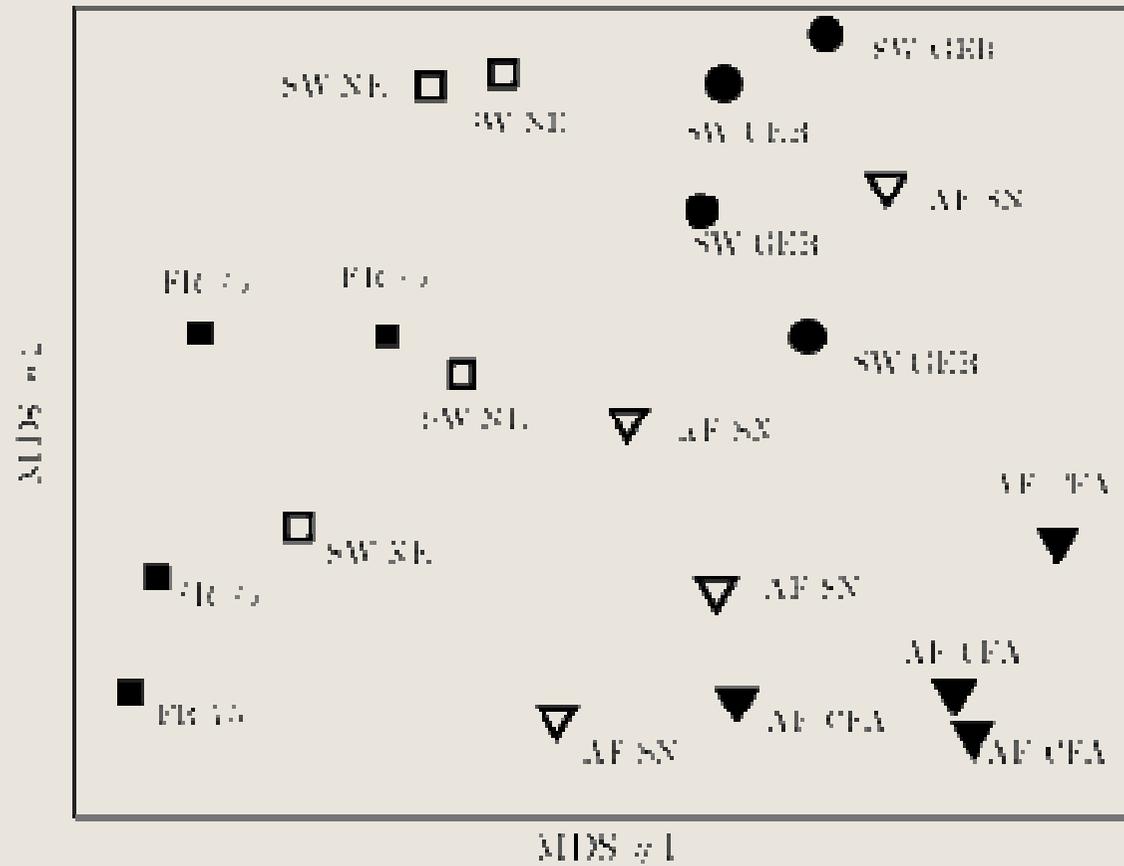
Durée des silences



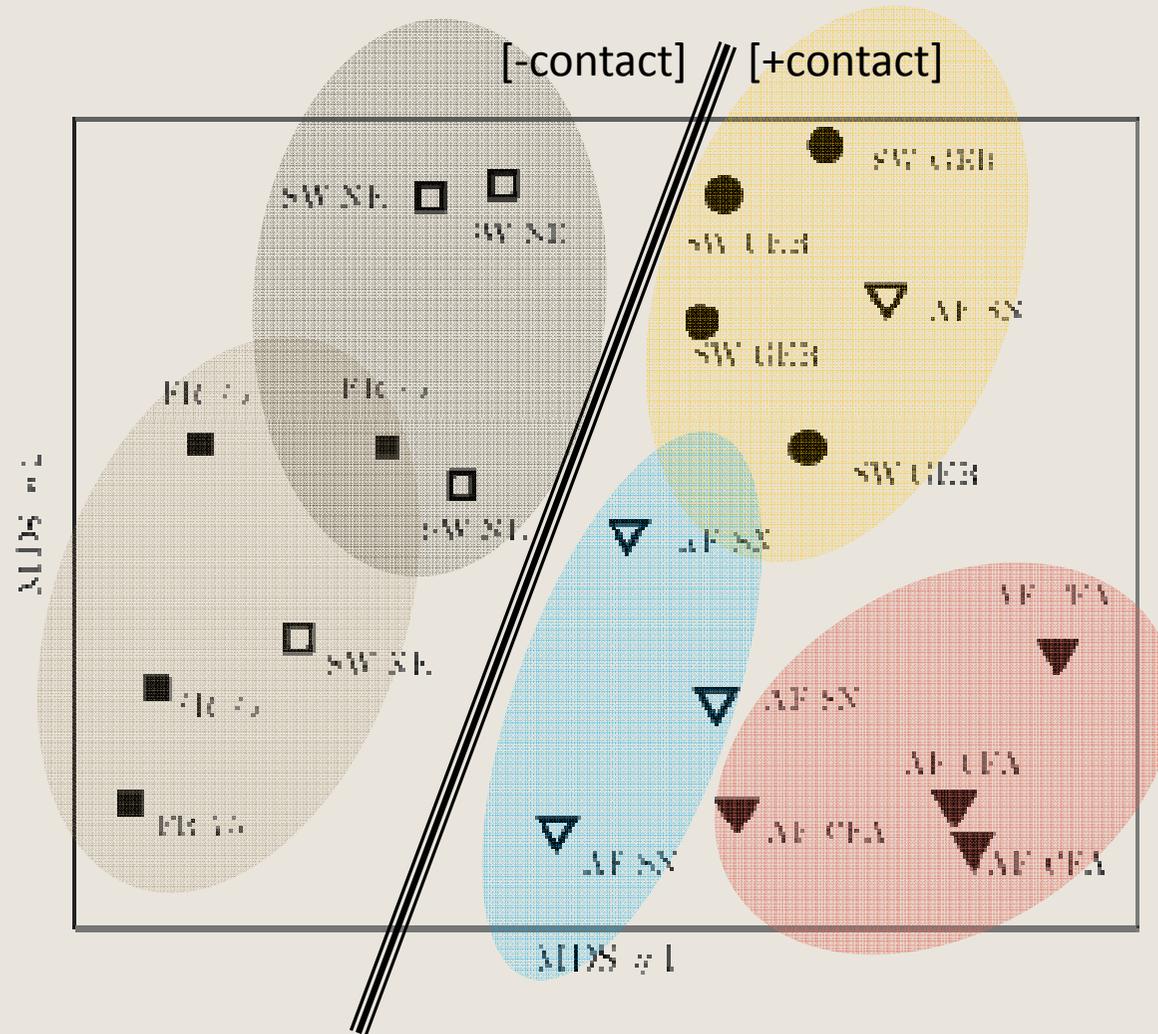
Synthèse

1. Prominences	
AP weight	+
Prominence ratio	+
2. AP restructuring	
Align-XP	-
Clash	++
3. Sandhis	
Sandhi_intem	++
Sandhi_extem	++
4. Rhythm	
Articulation rate	+
ΔC	+
%V	-
5. Silences	
Silence number	++
Silence duration (ms)	+

« Scaling » multidimensionnel



« Scaling » multidimensionnel



Conclusion

- Tous les paramètres étudiés ne se révèlent pas discriminatifs
 - Le respect ou non des règles de formation du SA comme « domaines » semblent les classifieurs les plus robustes
- Les locuteurs de L1 une langue [+ ton lexical] s'éloignent plus des variétés [- contact] que les locuteurs de L1 une langue [+ accent lexical]
- Les prédictions selon l'hypothèse de marque typologique se confirment globalement

Variations dans la formation du syntagme accentuel en français

Analyse de 5 points d'enquête PFC

Guri Bordal, Université Paris Ouest/Universitet i Oslo

Mathieu Avanzi, Université de Neuchâtel

Nicolas Obin, IRCAM/CNRS Paris

Alice Bardiaux, FNRS/Université catholique de Louvain