



# **Contribution de la prosodie à l'identification du parler de la région marseillaise**

Annelise Coquillon

## **► To cite this version:**

Annelise Coquillon. Contribution de la prosodie à l'identification du parler de la région marseillaise. 2004, pp.89-90. <hal-00135436>

**HAL Id: hal-00135436**

**<https://hal.science/hal-00135436v1>**

Submitted on 7 Mar 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# Contribution de la prosodie à l'identification du parler de la région marseillaise

*Annelise COQUILLON*

Laboratoire parole et langage, Institut de Phonétique  
Université de Provence, 29 Av. R. Schuman, 13621 Aix-en-Provence Cedex 1, France  
Tél. : +33 (0)4.42.95.36.27 - Fax : +33 (0)  
E-Mail : [annelise.coquillon@lpl.univ-aix.fr](mailto:annelise.coquillon@lpl.univ-aix.fr)

## ABSTRACT

This paper focuses on the identification of two different varieties of French on the basis of suprasegmental elements. More precisely, we will show that a perceptual discrimination between French spoken in the Marseille area and a more neutral French (which can be assimilated to standard French) is possible using prosodic cues alone, through low-pass filter speech. As the results show the importance of the suprasegmental elements of Marseille French, we investigated a preliminary research of the prosodic parameters of the corpus. We will show here that speech rate and global pitch register can discriminate the two varieties presented.

## 1. INTRODUCTION

L'étude d'une langue, et à fortiori de ses variétés, ne peut aujourd'hui se satisfaire d'une description purement segmentale, mais nécessite la prise en compte des phénomènes prosodiques, composantes essentielles dans la caractérisation linguistique des langues naturelles qui contribuent largement à différencier les langues entre elles. Une des questions en jeu ici est alors de savoir si cette fonction identificatrice de la prosodie joue un rôle pour caractériser les différentes variétés dialectales d'une même langue.

Nous nous attacherons ici à observer dans quelle mesure la prosodie permet en soi l'identification de variétés régionales et en l'occurrence de la variété de français parlée dans la région marseillaise.

## 2. MATERIEL

Dans le but de limiter les sources de variation dans notre corpus, nous avons choisi de comparer la variété marseillaise avec une variété de français plus "neutre" régionalement (groupe de contrôle). Cette démarche comparative permet la mise en relief des particularités dialectales de la région étudiée, en évitant l'introduction d'autres variables régionales. La région marseillaise est bornée par les anciennes limites linguistiques du provençal maritime, substrat du français parlé de nos jours dans cette région [5].

En ce qui concerne la nature du corpus, nous avons opté pour un corpus contrôlé, constitué de dialogues

interprétés par paires de locuteurs, qui permet une comparaison terme à terme des données.

L'origine géographique des locuteurs, ainsi que leur degré d'accent régional ont été évalués lors de tests de validation auditive (par un groupe de 35 sujets) qui ont permis de sélectionner cinq locuteurs marseillais et trois locuteurs non-méridionaux parmi les plus représentatifs de leur variété dialectale. Pour le groupe de contrôle, nous avons donc retenu les locuteurs originaires du Nord de la France et dont l'accent a été estimé comme n'ayant aucune marque régionale significative.

## 3. TEST PERCEPTIF

Afin d'évaluer l'importance relative de la prosodie dans l'identification de ces variétés régionales, nous avons élaboré un test de discrimination régionale, dans lequel le corpus a été modifié à l'aide d'une méthode classique de filtre passe-bas, qui élimine pour la majeure partie l'information segmentale (particularités phonémiques et sémantiques).

Les résultats de ce test, obtenus sur un groupe de 35 sujets d'origines géographiques diverses, se situent largement au delà du niveau de chance, la moyenne d'identification correcte de l'origine des locuteurs étant de 75,2 % (écart type 0,15). Ce score est proche de résultats obtenus dans des expériences similaires portant sur des langues différentes ([2], entre autres). Il permet de confirmer l'hypothèse selon laquelle une discrimination entre l'accent de la région marseillaise et un accent régionalement neutre est possible sur la base d'éléments prosodiques, et montrent l'importance de la prosodie dans la caractérisation régionale. Pour comparaison, les scores d'identification pour la même tâche en parole non-filtrée s'élevaient à 95,7 %.

## 4. ANALYSES PROSODIQUES

### 4.1 Vitesse d'élocution

Étant donné que certains auditeurs du test que nous venons de présenter affirmaient s'être basés, pour cette tâche de discrimination, sur le débit de parole des locuteurs, nous avons confronté les scores à ce paramètre, qui reste perceptible à travers le filtrage passe-bas. Nous avons alors observé que les locuteurs méridionaux présentent une vitesse d'élocution (5,79 syllabes par seconde) plus rapide que les non-

méridionaux (5,34 syll/sec). L'Anova révèle que cette différence est significative :  $F(1,6)=6,211$ ;  $p=0,047$ .

## 4.2 Registre tonal

Une première investigation sur les paramètres prosodiques globaux de notre corpus indique que le registre tonal, en termes de niveau et d'étendue [4], permet de distinguer les deux variétés régionales en question. Il est ici caractérisé par les valeurs des cibles tonales, qui correspondent aux changements de direction significatifs de la  $f_0$ . Elles ont fait l'objet d'une détection automatique (logiciel MOMEL) et d'un codage symbolique à l'aide de l'alphabet INTSINT [1]. En ce qui concerne le niveau, nous avons choisi de présenter ici, entre autres mesures, celles sur les moyennes de l'ensemble des points cibles, ainsi que sur les moyennes des cibles codées B (Bottom, valeurs minimales) et M (Mid, points moyens). L'étendue tonale est représentée ici, quant à elle, par la différence entre les moyennes des points absolus codés T (Top) et B. Les fréquences ont été normalisées en demi-tons.

**Tab. 1** – Niveau tonal (moyennes, points B et M) et étendue tonale (moyennes T/B) en demi-tons

locuteur	moy (sd)	B	M	T/B
M1	-0,63 (4,69)	-9,45	-1,71	17,85
M2	-0,57 (4,50)	-9,82	-4,14	17,75
M3	-0,47 (3,95)	-6,82	-3,48	15,36
M4	-0,47 (3,93)	-6,76	-2,49	17,22
M5	-0,39 (3,68)	-9,08	-2,88	16,54
NM1	-0,38 (3,56)	-6,44	-1,13	14,75
NM2	-0,34 (3,32)	-5,38	-0,19	13,35
NM3	-0,28 (3,06)	-5,57	-1,01	12,58

Pour chacune de ces mesures, nous pouvons observer que les locuteurs sont regroupés par région d'origine. Les Marseillais (M dans le tableau) présentent tous un niveau tonal en demi-tons plus bas que les non-méridionaux (NM). Ce facteur (origine régionale) a été validé statistiquement : L'Anova donne, pour les moyennes des fréquences :  $F(1,6)=8,401$  ;  $p=0,0274$  ; pour les moyennes des points cibles codés B :  $F(1,6)=8,043$  ;  $p=0,0297$  ; et M :  $F(1,6)=13,243$  ;  $p=0,0108$ . Nous pouvons également observer une tendance des locuteurs de la région marseillaise à utiliser une étendue tonale significativement plus importante que les non-méridionaux :  $F(1,6)=19,43$  ;  $p=0,0045$ .

Une corrélation de rangs de Spearman (niveau de significativité supérieur à 0,05) a démontré que ces mesures sur le registre étaient en étroite relation avec les scores d'identification du test de perception en parole filtrée (variable indépendante). Le niveau est inversement corrélé avec ce facteur :  $\rho$  (Rhô) = -0,905 pour les moyennes des points cibles ;  $\rho = -0,833$  pour les moyennes des B et  $\rho = -0,762$  pour celles des M. Enfin, l'étendue tonale (T/B) présente une corrélation positive de  $\rho = 0,929$  avec les scores d'identification régionale.

## 5. DISCUSSION ET PERSPECTIVES

Cette étude nous a permis de confirmer l'hypothèse selon laquelle une discrimination entre le français de la région marseillaise et une variété neutre est possible sur la base d'indices prosodiques. Outre la vitesse d'élocution et le registre tonal, nos perspectives de recherches pour la suite de ce travail consistent à approfondir la recherche des paramètres prosodiques pertinents dans cette discrimination. À un niveau plus local, nous avons identifié certains patrons tonals spécifiques au français parlé dans la région marseillaise. Nous avons également relevé des différences de comportement entre les deux variétés de notre corpus dans l'alignement de certains phénomènes tonals (pics de  $F_0$ , vallées) avec le segmental, notamment sur des items présentant un schwa final potentiel. Nous n'avons pas la place de détailler ici nos premiers résultats, mais il en ressort globalement que la caractérisation d'une variété régionale est plus le fait de l'interaction d'un ensemble de détails phonétiques (relevant tant du niveau segmental que prosodique) que la définition de marques régionales spécifiques et systématiques. Comme le montre [3], la pertinence des détails phonétiques est trop souvent négligée dans les descriptions linguistiques. Ainsi, si certains traits caractéristiques du français méridional peuvent se retrouver dans d'autres variétés, c'est non seulement leur fréquence d'emploi ou les modalités particulières pour lesquelles ils sont employés qui contribuent à caractériser cette variété, mais aussi et surtout, c'est la présence simultanée de ces différents indices qui fait la 'couleur' de cet accent.

## REFERENCES

- [1] D.J. Hirst, A. Di Cristo & R. Espesser. « Levels of representation and levels of analysis for the description of intonation systems » Dans M Horne: Prosody, Theory and Experiment, Dordrecht, Kluwer Academic Publishers, pages 51-88, 2000
- [2] M. Le Besnerais. « Reconnaissance par des locuteurs monolingues et bilingues de l'espagnol et du français à partir de productions filtrées » Actes des XXIe Journées d'Étude sur la Parole, pages 47-50, Avignon, Juin 1996
- [3] J. Local. « Variable domains and variable relevance : interpreting phonetic exponents. » Journal of Phonetics 31, pages 321-339, 2003
- [4] D. Patterson & D.R. Ladd. « Pitch range modeling: linguistic dimensions of variation » Actes du XIVe International Congress of Phonetic Sciences, pages 1169-1172, San Francisco, août 1999
- [5] P. Sauzet. « L'occitan : Langue immolée » Dans G. Vermes: Vingt-cinq communautés linguistiques de la France, vol. 1: Langues régionales et langues non territorialisées, Paris, L'Harmattan, pages 208-260, 1988