

# Role of segmental and suprasegmental cues in the perception of Maghrebian-accented French

Philippe Boula de Mareuil<sup>1</sup>, Belynda Brahimi<sup>1</sup> & Cédric Gendrot<sup>2</sup>

<sup>1</sup> LIMSI-CNRS, Orsay, France

<sup>2</sup> Laboratoire de Phonétique et Phonologie, Université de la Sorbonne nouvelle, Paris, France  
mareuil@limsi.fr

## Abstract

The general objective of this study is to clear up the relative importance of prosody in the identification of a foreign *accent*. The methodology we propose, based on the prosody transplantation paradigm, can be applied to different languages or language varieties. Here, it is applied to Maghrebian-accented French: we wanted to study what is perceived when the segmental and suprasegmental characteristics of Maghrebian and native French speakers are crossed. Results obtained with French listeners (*accent* degree rating task) and Algerian listeners (origin identification task) converge and suggest that the articulation of phonemes overrides prosody to account for Maghrebian *accents* in French.

## 1. Introduction

What is commonly meant by “foreign *accent*” results from the confrontation of two phonological systems in contact — from a mother tongue (L1) and a target language (L2). Beyond the relations which exist between these languages, a number of psycho- and socio-linguistic factors may come into consideration (even in the rating of a foreign *accent*) [1]. This article is devoted to Maghrebian-accented French (spoken by Arabic or Berber natives). Is it by chance that the term *accent* has been chosen to designate both a prosodic prominence and a foreign way of speaking? This question also applies to regional and social *accents* [2], and some authors such as Rossi seem to answer it in the negative [3]. Invariants probably exist, beyond regional differences, between a Parisian and a person from Marseilles speaking English. If the information carried by the segmental content (properties linked to the sound patterns) often enables us to identify a specific *accent*, prosody is also essential in language acquisition, and may trigger certain adjustments during second language learning. Although the characteristic categories of the mother tongue prosody are extracted very early by infants (prior to the lexicon acquisition [4]), prosody is one of the most ticklish elements to teach in a second language.

If phonemic characteristics and prosodic dimensions are important, a foreign *accent* should reflect both segmental and suprasegmental difficulties. Recent studies thus approached the contribution of prosody to the perception of a foreign *accent* [5; 6]. Likewise, some investigations brought to light the important or even major role of prosody in the identification of regional *accents* or dialectal *varieties* [7; 8]. The TIE network (Tone and Intonation in Europe) [9] in particular aims at providing a platform to coordinate these prosodic analyses of European languages and non-standard varieties (including immigrant communities) in the laboratory phonology approach. A number of studies especially within

the framework of the IViE project (Intonational Variation in English) revealed how certain intonational patterns are specific to some English varieties [10]. The AMPER [11] linguistic atlas of Romance languages considers French among others: it is a continuation of a series of inquiries concerning different stress patterns in *patois* from the North/North-East of France, and more generally from the whole French-speaking area [12; 13]. To complete the survey, let us cite a few additional studies on the identification of Arabic dialects [14; 15]. The latter study [15] was carried out using procedures put forth in Ramus’s dissertation [4] on the different rhythmic classes — especially syllable-timed *vs.* stress-timed. Others attempt to elucidate the role of intonation rather than rhythm (TIE and IViE projects in particular). However, Arabic and Berber prosody has barely been tackled, and remains somewhat of an unexplored field — not to mention analyses of Arabs and Berbers speaking French. With over 900,000 adult speakers living in Metropolitan France, Arabic is nonetheless the main language other than French spoken in France, before Alsatian (600,000 speakers), according to a recent inquiry [16]. For a discussion on Berber and in particular on Kabyle (Algerian Berber), one can refer to [17]. For Arabic and in particular Moroccan Arabic, see [18].

In our study, the first objective of which is to appreciate what constitutes the Maghrebian *accent*, 20 or so native speakers of Arabic, Kabyle and French were recorded. A perceptual study was conducted to sort out what is perceived when the segmental and suprasegmental features of the different speakers are crossed. The proposed experiment could be replicated with other L1-L2 pairs, and may find applications in automatic language/dialect identification. It is described in next section, with an *accent* degree evaluation task performed by French listeners and an origin identification task performed by Algerian listeners. Results are provided in section 3: they suggest that the articulation of phonemes and voice quality take the first place as compared to prosody, to identify a Maghrebian *accent* in French.

## 2. Experiment

### 2.1. Material

We instructed 18 speakers to read a 400 word text: 6 Arab (2 Moroccan, 2 Algerian and 2 Tunisian), 6 Kabyle and 6 French speakers living in the Paris region. The subjects were 25 years old on average, and the Maghrebians had arrived in France for 2 years on average. Six other speakers of these different origins were recorded, for a learning phase. All of them were students or young doctors, and there were as many males as females in each group. The speakers had no precise idea on the present study. They were requested to read the text

of the PFC project (Phonology of Contemporary French) [19], which was split into sentences: this way, if they came up against a sentence or hesitated, they would resume the reading from the beginning of the sentence. The recordings took place in Orsay, in a soundproof booth with a high-quality microphone located about 20 cm from the mouth, and were digitised at 22.05 kHz, 16 bits, mono.

We extracted a dozen sentence pairs (the ones stemming from Maghrebian speakers, the others from French speakers or Maghrebian speakers of different origins). Energy was normalised, and the segmental and suprasegmental characteristics of the different populations were crossed. The sentences were selected on the basis of 3 criteria:

- presence of presumed typical (supra)segmental feature, as in previous studies [5];
- matching realisations of liaisons and schwas;
- matching realisation of pauses and pitch — even if we can play on the latter parameters.

The sentences, which ranged from 2.4 to 15.7 seconds, were 8.7 seconds long on an average.

## 2.2. Prosody transplantation: methodology

In order to copy the prosodic parameters from a speaker to another one, a script was written for the PRAAT software (<http://www.fon.hum.uva.nl/praat/>), which enables signal handling and re-synthesis with the help of the PSOLA algorithm. It follows the upcoming scheme: extraction and then transplantation, phoneme by phoneme, of durations and then pitch.

Before starting on this phase, a segmentation into phonemes is required: it was made by an expert, under PRAAT, on the basis of perceptual and visual clues — on spectrograms. The analysis of phoneme durations, pauses and fundamental frequency ( $F_0$ ) is then automatic. After possible hand-corrections, the transplantation of these parameters amounts to:

- checking that the number of segments in the two sentences of the chosen pair are equated;
- calculating lengthening or shortening coefficients for each phoneme or pause of a speaker with respect to another one;
- for each phoneme or pause, building new durations which will replace the original durations of either sentence;
- grafting  $F_0$  from one of the sentences onto the other one, and vice versa.

Two new stimuli are thus obtained: the block diagram of Fig. 1 displays the different steps of the process.

## 2.3. Perceptual tests

In order to disentangle the influence of segmental vs. suprasegmental feature in the perception of Maghrebian-accented French, two types of perceptual tests were run: the first one concerns French listeners (with an *accent* degree

estimation task); the second one involves Algerian listeners (with an origin identification task). The choice in favour of this second population is dictated by their geographical position in the heart of Maghreb and their familiarity with Kabyle. Not counting a learning phase (or more exactly of familiarisation with the type of stimuli), each test was composed of about 40 stimuli, whether original or resulting from the different prosody transplantations. Each experiment lasted about 15 minutes. The listeners, all with normal hearing, were not specialist in linguistics, and were not paid for this task. They were not urged to answer; but they could listen to each stimulus only once. The tests were conducted in the Paris region, in a quiet room, over headphones, through the E-prime software, which offers a user-friendly interface to read the instructions, listen to the stimuli and capture the answers.

Firstly, native French listeners were asked to rate the degree of accentedness on a 9-point scale (from 0 to 8). 0 represents standard Parisian French, without any marked accent; 1 designates a very mild accent (only punctually marked); 2 a mild accent; 3 a moderate accent; 4 an average accent; 5 a rather strong accent; 6 a strong accent; 7 a very strong accent; and mark 8 designates so strong an accent that it poses intelligibility problems. The learning corpus is composed of 10 stimuli, providing landmarks to assess the degree of accentedness of male and female natural voices as well as prosody modifications not used in the sequel. As for the test, it comprises 42 randomised stimuli (18 originals + 24 French-Maghrebian crossings).

Secondly, 15 native Algerian listeners were asked to identify the origin (Kabyle, Moroccan, Algerian or Tunisian Arab) of the speakers they would hear. The listeners, who had arrived in France for 3 years on average, did not know the used voices. The learning corpus is composed of 4 stimuli, giving a sample of male and female voices of Kabyle, Moroccan, Algerian or Tunisian Arabian origin, as well as prosody transplantations not used in the remainder of the trial. As for the test, it comprises 40 randomised stimuli (12 Maghrebian originals + 12 Maghrebian-French crossings + 16 inter-Maghrebian crossings).

## 3. Results

### 3.1. Test with French listeners

The test conducted with French listeners shows that the Maghrebian *accent*, rated as average (about 4 out of 8) is judged as stronger after prosody modification: this artifact is due to the acoustic degradation, which listeners judge even though their orders were not to concentrate on it. The weak degree (<1) attributed to the original French speakers is reassuring as for the methodology. But above all, if we compare all other things being equal, a Maghrebian voice with a French prosody is judged as having a stronger *accent* degree than a French voice with a Maghrebian prosody. The difference is highly significant according to a paired *t*-test ( $df=179$ ;  $p<0.01$ ), and it is more important than between original/modified Maghrebian voices (see Table 1).

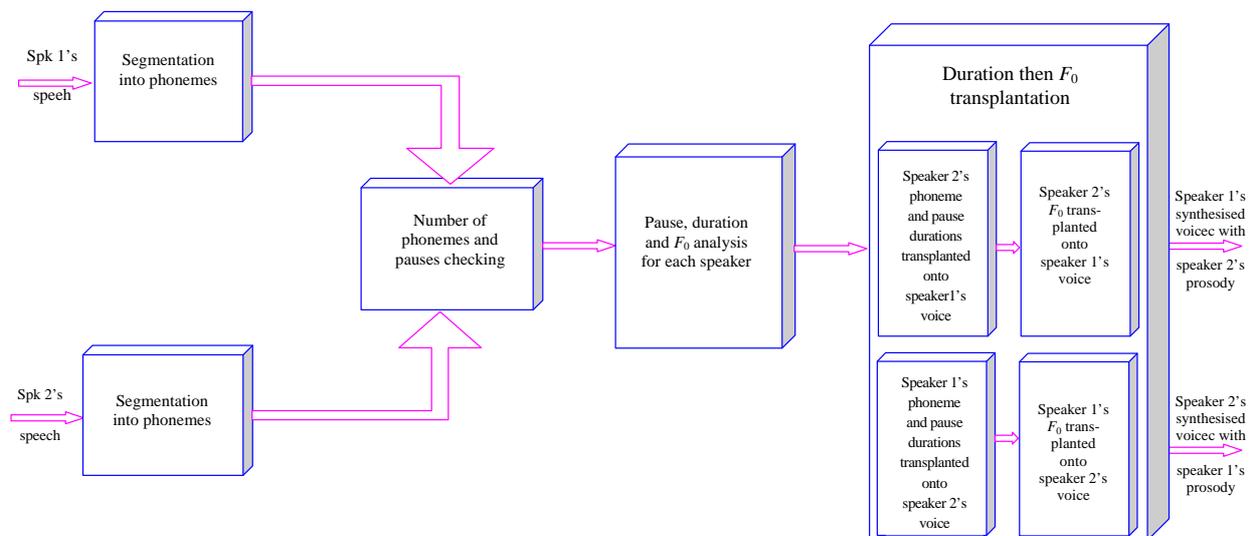


Figure 1: Prosody transplantation.

Table 1: Degree of accentedness judged by 15 French listeners, on a 0-8 scale.

original French voice	0.8
French voice, Maghrebian prosody	3.2
Maghrebian voice, French prosody	4.4
original Maghrebian voice	3.9

This tendency in favour of a greater role played by segmental cues is illustrated by the Maghrebians' tendency to pronounce [i] for /e/, for example in the word *télévision*. Formant measures confirm this fact (with F1 values in the order of 300 Hz for Algerians). Instead, if we compare  $F_0$  targets and phoneme durations, we can see that these ones are very close even between sentences which sounded to indicate a reliable effect of prosody.

If segmentals override suprasegmentals in what is perceived as a Maghrebian *accent*, the question which naturally arises is to know to what extent the articulation of phonemes enables us to identify the origin within Maghreb of this *accent*. Preliminary tests showed that French listeners albeit familiarised with the Maghrebian *accent* found it hard to distinguish Arabian and Kabyle *accents*: hence the follow-up test.

### 3.2. Test with Algerian listeners

The results of the test conducted with Maghrebian listeners reveal the difficulty in identifying the origin of the stimuli taken as a whole, even for Algerians (see Table 2). This is at variance with the listeners' impression that they can easily recognise a Kabyle speaker, for instance, whose L1 is not Arabic. Now the Kabyle origin is not better recognised than chance level (25%). Only the Moroccan origin (by more than 56%) is well recognised, and highly significantly better than

chance (according to a  $\chi^2$  test,  $df=3$ ,  $p<0.01$ ). As for the speakers' origin with respect to their prosody, it is recognised even worse.

Table 2: Identification scores of the origin of 12 Maghrebian speakers with respect to their segmentals — in row the origin of voices, in column responses (%).

voice \ answer	Moroccan	Algerian Ar.	Kabyle	Tunisian
Moroccan	<b>56.2</b>	13.3	12.4	18.1
Algerian Ar.	<b>32.5</b>	20.0	20.0	27.5
Kabyle	21.8	26.3	<b>28.1</b>	23.7
Tunisian	16.2	<b>41.9</b>	11.4	30.5

A principal component analysis was performed, so as to determine the influence of sentence length and *accent* degree (rated by the 15 French listeners) on the Algerian listeners' judgements, in terms of correct answers on the voices of various origins. The following  $r$  correlation coefficients were found: respectively -0.15 (no correlation) and 0.40 (weak correlation, as confirmed by a comparison with Student's  $t$ -distribution tables [ $t_{0.95}>1.39$ ;  $p>0.05$ ]).

Table 3: Identification scores of the origin of the Maghrebian speakers having a strong *accent* with respect to their segmentals. Percentages are given with respect to 15 listeners  $\times$  4 types of prosody per voice.

voice \ answer	Moroccan	Algerian Ar.	Kabyle	Tunisian
Moroccan	<b>63.3</b>	5.0	10.0	21.7
Algerian Ar.	18.3	25.0	11.6	<b>45.0</b>
Kabyle	18.3	10.0	<b>60.0</b>	11.6
Tunisian	22.7	28.3	11.6	<b>38.3</b>

If however we extract the speakers of each origin judged as having the strongest *accent* according to the French listeners, we can observe better origin identification scores — even if these ones are established from smaller data. The *accent* degrees are 4.2 for the Moroccan speaker, 4.0 for the Algerian Arab speaker, 7.0 for the Kabyle speaker and 4.5 for the Tunisian speaker. Results on the basis of their sole voices (their segmentals, whatever the prosody) are very revealing (see Table 3): the identification scores of the Moroccan and Kabyle origins are highly significantly above chance (according to a  $\chi^2$  test,  $df=3$ ,  $p<0.01$ ). The Tunisian origin is also recognised with scores significantly above chance (according to a  $\chi^2$  test,  $df=3$ ,  $p<0.05$ ). Only the Algerian Arabian origin is poorly recognised.

By and far, also, the prosody of these speakers having a rather strong *accent*, when applied to various voices, is not so well recognised in the confusion matrix of Table 4 below as it is in Table 3. The same happens with Kabyle, which again suggests that the very strong *accent* degree (7 out of 8) would be due to segmentals rather than suprasegmentals. The Moroccan and Tunisian origins are rather well identified, but the results are not (highly) significant. The Algerian Arabian origin, for its part, remains badly recognised.

Table 4: Identification scores of the origin of the Maghrebianspeakers having a strong *accent* with respect to their prosody. Percentages are given with respect to 15 listeners  $\times$  3 voices per type of prosody.

pros. \ answer	Moroccan	Algerian Ar.	Kabyle	Tunisian
Moroccan	<b>40.0</b>	15.6	31.1	13.3
Algerian Ar.	13.3	37.8	8.9	<b>40.0</b>
Kabyle	<b>37.7</b>	15.6	35.6	11.1
Tunisian	28.9	17.8	15.6	<b>37.8</b>

#### 4. Conclusion

At the term of this analysis and given the achieved results, segmental factors seem to be more relevant than prosody in the perception of Maghrebianspeakers. A fortiori, segmental parameters take the first place when speakers have a strong *accent*. Nevertheless, the *accent* origin identification, except for the strongest *accents*, proves delicate: this is in keeping with earlier studies on the identification of Maghrebianspeakers, within which a considerable variability is prevalent [14]. Interestingly however, the Moroccan *accent* origin is honourably recognised: better than the Kabyle origin, despite a lower *accent* degree, and although Kabyle speakers have a mother tongue which distinguishes them from other Maghrebianspeakers, of Arabic L1. The question this raises deserves to be addressed more thoroughly, as well as the perception of Maghrebianspeakers vs. Middle Eastern *accents* in French, with more speakers and listeners. Finally, results might be different on no longer read but spontaneous speech: it is fitting to remain cautious.

#### 5. References

[1] Flege, J.E., Munro, M.J. & MacKay, I.R.A., “Effects of age of second-language learning on the production of English consonants”, *Speech Communication*, 16:1-26, 1995.  
 [2] Léon, P., *Précis de phonostylistique. Parole et*

*expressivité*. Fernand Nathan, Paris, 1993.  
 [3] Rossi, M., *L’intonation, le système du français : description et modélisation*, Ophrys, Paris, 1998.  
 [4] Ramus, F., *Rythme des langues et acquisition du langage*, PhD thesis, EHESS, Paris, 1999.  
 [5] Jilka, M., *The contribution of intonation to the perception of foreign accent. Identifying intonational deviations by means of F0 generation and resynthesis*, PhD thesis, University of Stuttgart, 2000.  
 [6] Boula de Mareüill, P., Marotta, G. & Adda-Decker, M., “Contribution of prosody to the perception of Spanish/Italian accents”, *2<sup>nd</sup> International Conference on Speech Prosody*, pp. 681-684, Nara, 2004.  
 [7] Coquillon, A., Di Cristo, A. & Pitermann, M., “Marseillais et Toulousains gèrent-ils différemment leurs pieds ? Caractéristiques prosodiques du schwa dans les parlers méridionaux”, *XXIII<sup>es</sup> Journées d’Étude sur la Parole*, pp. 89-92, Aussois, 2000.  
 [8] van Leyden, K. & van Heuven, V.J., “Prosody versus segments in the identification of Orkney and Shetland dialects”, *15<sup>th</sup> International Congress of Phonetic Sciences*, pp. 1197-1200, Barcelona, 2003.  
 [9] Gussenhoven, C., *The Phonology of Tone and Intonation*, in preparation.  
 [10] Grabe, M.E., Post, B. & Nolan, W.F., “Modelling intonational variation in English: the IViE system”, *Prosody 2000*, pp. 51-57, Kraków, 2000.  
 [11] Contini, M. et al., “Un projet d’Atlas Multimédia Prosodique de l’Espace Roman”, *1<sup>st</sup> International Conference on Speech Prosody*, pp. 227-230, Aix-en-Provence, 2002.  
 [12] Carton, F. “Insistance dialectale”, *Studia phonetica*, 12:59-92, 1977.  
 [13] Hintze, M.-A., Pooley, T. & Judge, A. (eds), *French accents: phonological and sociolinguistic perspectives*, AFLS/CiLT, London, 2000.  
 [14] Barkat, M., *Détermination d’indices acoustiques robustes pour l’identification automatique des parlers arabes*, PhD thesis, Université Lumière, Lyon 2, 2000.  
 [15] Hamdi, R., “Variation rythmique des dialectes arabes”, *XXIV<sup>es</sup> Journées d’Étude sur la Parole*, pp. 68-71, Nancy, 2002.  
 [16] Héran, F., Fillon, A. & Deprez, C., “La dynamique des langues en France au fil du XX<sup>e</sup> siècle”, *Population et Sociétés*, 376:1-4, 2002.  
 [17] Chaker, S., “Données exploratoires en prosodie berbère : I. L’accent kabyle”, *Comptes Rendus du GLECS*, 31:27-54, 1987.  
 [18] Bnkirane, T., “Intonation in Western Arabic”, In D. Hirst & A. Di Cristo (eds.), *Intonation Systems. A Survey of Twenty Languages*, Cambridge University Press, Cambridge, 1998.  
 [19] Delais-Roussarie, É. & Durand, J., *Corpus et variation en phonologie du français : méthodes et analyses*, Presses Universitaires du Mirail, Toulouse, 2003.