Corsican French questions: is there a prosodic transfer from Corsican to French and how to highlight it?

Philippe Boula de Mareüil[1], Albert Rilliard[1], Paolo Mairano [2], Jean-Pierre Lai [2]

[1] LIMSI-CNRS, Orsay, France
[2] GIPSA-lab, Grenoble, France

philippe.boula.de.mareuil@limsi.fr, albert.rilliard@limsi.fr,
paolomairano@gmail.com, Jean-Pierre.Lai@u-grenoble3.fr

Abstract

This study investigates whether a prosodic transfer can be highlighted from Corsican (an Italo-Romance language) to French spoken in Corsica, where French is now the dominant language. A corpus of transparent sentences such as *la touriste trouve la caserne* (French) or *a turista trova a caserna* (Corsican) was designed and the productions of bilingual speakers, recorded in Corsica, were compared with the French counterparts of Parisian reference speakers. The melody of yes/no questions turns out to contrast Corsican and Corsican French (both with high tones followed by final pitch falls) and standard French (with utterance-final high tones). The former pattern can be interpreted as a prosodic transfer from Corsican to French. Various methods are considered to validate this hypothesis and an experimental paradigm is proposed.

Index Terms: prosody in contact, questions, Corsican accent in French, endangered languages

1. Introduction

This article explores the issue of prosodic transfers in languages in contact. It presents a case study (Corsican French) and offers a review of methods suited to support prosodic transfer hypotheses.

On the Mediterranean island of Corsica (whose population is close to 300,000), French has become the first language, before Corsican (an Italo-Romance language of the Tuscan group). Corsican has been removed from the Italian sphere of influence since the attachment of Corsica to France in 1768-1769. This *Ausbau* language in Kloss's terminology [1] (i.e. a dialect which has reached the dignity of language) is a *polynomic* language. This concept was developed by Marcellesi [2] to account for the dialectal diversity of languages which are variation-tolerant. In Corsica, a broad North/South division is noticeable, with southern varieties of Corsican being most conservative [3][4]: they are close to the Corsican dialect spoken in Gallura (northeastern Sardinia).

In Corsica (in contrast to Sardinia), Corsican cohabits with various forms of the French language: an academic and official (Parisian) variety which reflects an idealised conception of language, a Corsican French variety which is no more than the standard (or standardised) variety spoken with a Corsican accent, several imported varieties (which are approached as Parisian French, southern French or *Pied-Noir* — North-African — French), slang broadcast by the media, and a hybrid dialect built on a Corsican substratum. This so-called Francorsican variety is a substitute for the regularly-declining Corsican language [3]. Its most often reported elements are more or less stable lexical items such as *stamper* ('copier' < *stampà* 'to print'), *strapper* ('to tear' < *strappà* 'to break'). Code-switching and peculiar grammatical constructions also receive some attention, but very few studies account for Corsicans' pronunciation in French, even though a Corsican accent is often parodied by humourists.

The Corsican-characteristic lenition of some consonants called *cambiarine* ('mutter consonants') may be observed in Corsican French, even though it is much rarer than in Corsican. In Corsican, for instance, /k/ weakens into [g] and /g/ is elided in a number of intervocalic contexts. As for prosody, some rising-falling melodic clichés in interjections or vocatives (e.g. *o Francè* 'hey Franck') may be shared by Corsican and Corsican French [5]. Corsicans' low speech rate, often used by humourists, rather results from caricature.

This paper reports on a first analysis of data collected during fieldwork in Corsica compared with Parisian reference speakers. It is organised in the following way: the next section (section 2) presents the survey and the corpus. Four Corsican-French bilingual speakers were selected and four standard French speakers were then recorded. Section 3 provides a descriptive analysis of prosody in Corsican and French questions, in which a "Corsican accent" is perceptible in Corsicans speaking French. Section 4 then addresses methodological issues to evidence a possible prosodic transfer. Various techniques to tease apart prosody are reviewed, and a new experimental setup is proposed, whose results will hopefully be available at the time of the conference. Section 5 is a tentative conclusion.

2. Survey and corpus

Recordings in Corsican and French were made around Corti, in the centre of Corsica. Corti was the former capital of the independent Corsica (1755–1769). It also hosts the University of Corsica which was founded at that time and was reopened in 1981. In this regard, it is known as a hotbed of Corsican militancy. Our quest for proficient Corsican speakers led us to pursue our investigation in the nearby villages of Loretu di Casinca and Pedicorti di Gaghju, on the outskirts of the Castagniccia region.

2.1. Material

In total, seven Corsican speakers were recorded (with a high-quality device, at 44.1 kHz):

- uttering 60 or so sentences with very controlled structures, repeated in declarative and interrogative modalities (designed in such a way as to be relatively transparent in Corsican and French);
- reading the French version of the fable "The North Wind and the Sun" and translating it into Corsican;
- in semi-directed interviews in both French and Corsican.

For most speakers, maptask interactions were also recorded.

Table 1. *Examples of sentences in Corsican and French (with the English translations).*

Corsican	French	English
A turista trova a cavità prufonda.	La touriste trouve la cavité profonde.	The tourist finds the deep cavity.
U pudestà malatu trova a caserna.	Le podestat malade trouve la caserne.	The sick *Podestà* finds the barracks.
A femina di l'avviò trova u limitu.	La gamine de l'avion trouve la limite.	The girl of the plane finds the limit.

The controlled sentences (presented in random order in the form of drawings with legends) met the requirements of the AMPER (Multimedia Prosodic Atlas of the Romance Area) project [6], as one of the aims of our fieldwork was to enrich this dialectological atlas and to allow comparisons with other Romance dialects (especially from Sardinia). In compliance with the AMPER protocol, the designed sentences (12-syllable long, on average) needed have dissyllabic verbs, trisyllabic nouns and expansions with various accentual patterns. Examples of such sentences are displayed in Table 1.

In French, stress always falls on the last syllable of word phrases (or on the syllable preceding a pronounced word-final schwa). By contrast, in Corsican, words may be oxytone, paroxytone or proparoxytone (i.e. stressed on the last, penultimate or antepenultimate syllable, respectively). Trisyllabic adjectives cannot be oxytone in Corsican: we thus used prepositional phrases such as *di l'avviò* (Fr. *de l'avion*, 'of the plane'), which is a French borrowing. In addition, we ensured that the French counterparts would be as close as possible to Corsican. We included as many consonant clusters as possible to increase the chance that final schwas would be pronounced. The final schwa which is often dropped in non-southern French is most likely to be pronounced when it is surrounded by at least three consonants [7], as is the case in *la touriste trouve* ('the tourist finds').

For oxytone words, we selected concrete nouns such as *cavità* (Fr. *cavité*, 'cavity') and *pudestà* (Fr. *podestat*, '*Podestà*'). Since the latter is masculine (contrary to the former), the adjective which could accompany it had to have the same form in the masculine and in the feminine in French, to keep the number of syllables unchanged. We selected adjectives such as *bulgaru/a* (Fr. *bulgare*, 'Bulgarian').

For paroxytone words, we selected nouns such as *caserna* (Fr. *caserne*, 'barracks'). For proparoxytone words, we selected nouns such as *limitu* (Fr. *limite*, 'limit'), which is feminine in French.

2.2. Selected speakers and sentences

Four bilinguals, very committed in the cultural and linguistic field, were selected for this study: two males (aged 35 and 57) and two females (aged 50 and 72). They were compared with Parisian speakers, matched in age and gender, who were asked to pronounce the same list of sentences — in French.

As could be expected, no stress pattern transfer to French was observed for Corsican proparoxytone words such as *pubblica* (Fr. *publique*, 'public'). Nevertheless, similarities are noticeable between the pitch contours of Corsican paroxytone words and their French counterparts. A Corsican accent is especially perceived in questions. In view of perceptual experiments, we selected seven yes/no questions of the type "a turista trova a caserna?" (Fr. "la touriste trouve la caserne?", 'the tourist finds the barracks?').

3. Prosody of Corsican/French questions

In the bilinguals' utterances, whether in Corsican or French, a high tone is noticeable at the beginning of questions (with no precise syllabic anchoring), whereas the utterance-final stressed syllable is realised with a pitch fall. This is particularly striking in the pitch curve of Corsican (see Figure 1a), in which most parts are voiced due to the lenition phenomena described above. The pitch curves of the corresponding sentences produced by the same speaker in French and another bilingual speaker in Corsican are shown in Figures 1b and 1c. In comparison, the pitch curve of a Parisian speaker for the same French sentence exhibits a sharp pitch rise at the end of the question (see Figure 1d). The Praat software (www.praat.org) was used, with manual corrections, to extract fundamental frequency (F_0).
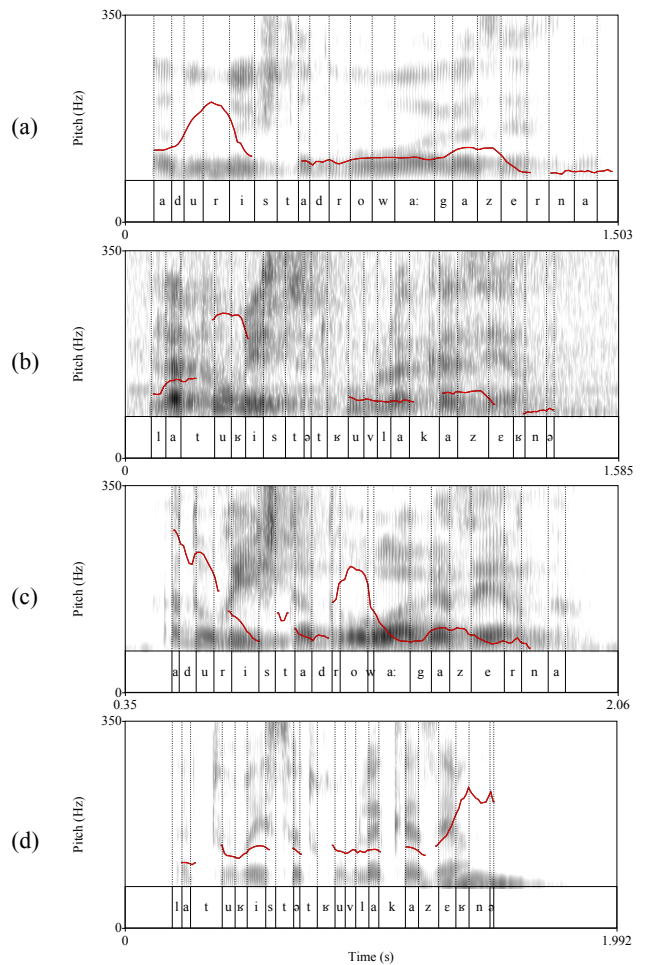

Figure 1: *Spectrogram and pitch curve of the sentence 'the tourist finds the barracks?' uttered (a) in Corsican by a bilingual male speaker, (b) in French by the same Corsican speaker, (c) in Corsican by another bilingual male speaker, (d) in French by a Parisian male speaker.*

To quantify these prosodic tendencies, the speakers' questions were hand-segmented into vowel nuclei. F_0

measurements were accordingly taken (using Praat) at the beginning, the midpoint and the end of each vowel. After manual correction, the results were plotted as exemplified in figure 2 for the F_0 representation of a sentence in Corsican, Corsican French and Parisian French.
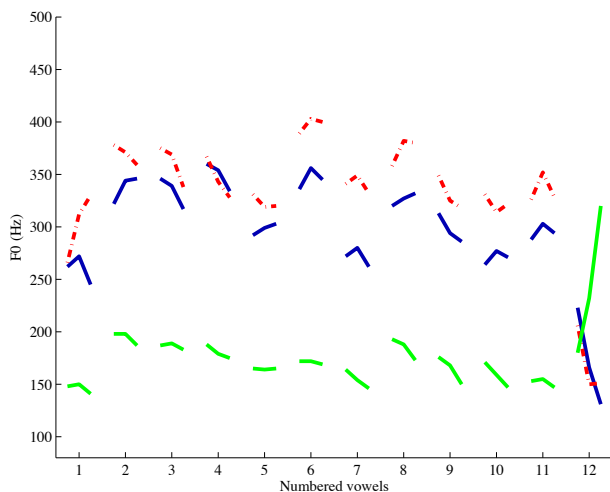

Figure 2: *duration-independent vowel-based F_0 representation of the sentence "the sick Podestà finds the cavity?", uttered by a female speaker in Corsican (dark blue), the same female speaker in Corsican French (dashed-dotted red) and by another female speaker in Parisian French (light green). The cross-speaker pitch range difference is a coincidence.*

In almost all cases, the F_0 peak is located in the end of the question (the verb phrase) in Parisian French: on the penultimate or on the final vowel of the utterance. By contrast, in the majority of cases, the maximum F_0 value is located in the beginning of the question in Corsican and Corsican French. Most exceptions, both in Corsican and Corsican French, come from one of the Corsican females (a trainer for bilingual school teachers, aged 50) who in our perception has a mild Corsican accent when speaking French. This is in line with sociolinguistic studies according to which a regional accent is often considered as an attribute of masculinity [8][9]. In Corsican too, this 50 year-old female speaker exhibits intonational patterns which are closer to French. Of course, the prosodic transfer from Corsican to French is not systematic. Yet, interesting trends show up in figure 3, considering initial F_0 peaks as bearing on one of the first four vowels (i.e. on the subject noun phrase) and final peaks as bearing on the last or penultimate vowels.
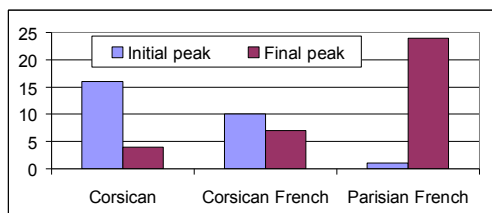

Figure 3: *Initial and final F_0 peaks in Corsican, Corsican French and Parisian French questions. The values indicated by the bars do not add up to a constant value for each variety because some F_0 maxima are utterance-medial (e.g. on the verb).*

Another way of quantifying differences across Corsican and French varieties consists of calculating pitch difference between the midpoints of the last stressed vowel of each question (bearing the so-called nuclear accent) and the vowel preceding it. Mean values in semitones (ST) are -3 ST for Corsican, -2 ST for Corsican French (both corresponding to falling slopes) and 4 ST for Parisian French (corresponding to a rising slope).

4. Methodological issues to sort out the contribution of prosody

If the pitch differences mentioned above are meaningful, they should be relevant in perception. This section is an overview of a few studies which attempted to elucidate the role of prosody in accented speech, with particular attention paid to the French language.

Previous studies have concentrated on prosody in a second language (L2) [11][12][13][14] or languages in contact [15]. The former studies used prosody manipulation, modification and synthesis, as others did in order to investigate prosody in several dialectal varieties [15][16]. In [17], two techniques were utilised:

- based on diphone synthesis (as in [12]);
- modifying the prosody of the natural voice (as in [13]), by using the PSOLA speech processing algorithm implemented in Praat.

4.1. Applications to French

The two approaches were applied to French. In particular, [18] used the second approach to clear up the relative importance of prosody in the identification of Maghrebian-accented French. However, results obtained with native French and Algerian listeners showed a minor role of prosody.

In [19], these two approaches were combined by grafting the duration and F_0 parameters onto the male and female voices of a unit-selection speech synthesis system (www.acapela-group.com). The original speech was segmented and labelled into phonemes. The phoneme string was then given as input to the speech synthesis system, and the synthesis output was segmented in the same way. Phoneme by phoneme, the duration and F_0 parameters of the originals were then transplanted onto synthesis voices. Speech samples from Polish learners, late and early bilinguals were examined and compared to those of Parisian French monolinguals. Their native-likeness was assessed by experts in prosody on both the original signal and synthesis voices using prosody transplantation. However, results on prosody transplantation only showed a significant effect for the group of learners: prosody solely enabled proficient speakers to be distinguished from learners.

When a local prosodic phenomenon can be identified as typical of a given language variety, it is easier to test its perceptual relevance. For example, sociophonetic work suggested that an atypical prosodic pattern, namely a word-final sharp pitch fall, could characterise the French vernacular of youth living in working-class suburbs (the so-called "suburban accent"). A question investigated in [20] is whether the presence/absence of these prosodic patterns increases/ decreases the perception of the suburban accent. Using prosody modification and resynthesis, perceptual experiments were conducted. Results involving listeners from two French regions showed that utterances with (respectively without)

high-low pitch falls are perceived as presenting a higher (resp. lower) degree of accentedness.

The same method, with an accent rating task, could be applied to Corsican French. However, to explicitly grasp a possible prosodic transfer, we found it more interesting to use delexicalised speech, with an ABX protocol. The next sub-section describes the experiment we endeavoured to conduct.

4.2. Proposed experiment on Corsican French

A test was designed with an ABX task consisting of asking subjects whether the intonation of stimulus X is closer to A or B. A and B stimuli are standard French and Corsican sentences which are the translations of one another, taken from age-matched speakers of the same gender. X is the corresponding sentence pronounced in French by a Corsican or a Parisian speaker (different from A and B but of the same gender).

Various delexicalisation techniques were considered, including low-pass filtered speech as used by [11]. Yet, low-pass filtered speech at 400 Hz preserves enough information to enable the recognition of the language (French or Corsican). The bias this would have induced led us to opt for other delexicalisation procedures such as the ones proposed by [21], based on text-to-speech synthesis. Finally, the *hum*-based resynthesis of the pulse train implemented in Praat was judged to yield the most ecological speech material. This material is currently being used in our ABX perception test, with A and B stimuli being delexicalised.

5. Conclusions

This article presented fieldwork carried out in Corsica. The prosodic structures of transparent sentences in French and Corsican were compared. In particular, yes/no questions were analysed. A high tone at the beginning of the question and a pitch fall at the end were observed in both Corsican and Corsican French, and contrast with the melody of standard French questions. The same prosodic shape (with a pitch peak enabling distinctions between questions and statements) was found in Sardinian and in regional Italian as spoken in northern Sardinia [14]. Yes/no questions with terminal falling pitch contours have also been reported (and analysed within the framework of Intonational Phonology) in southern varieties of Italian [22][23]. A common substratum may explain this phenomenon. Yet, an interpretation in terms of prosodic transfer still deserves validation throughout percep-tual experiments and an examination of spontaneous speech.

Speech synthesis has been used in previous work to disen-tangle prosody and segmentals. It will be applied to French in contact with Corsican in a near future. A comparison with Occitan dialects spoken in the South of France is also in progress. The same methodology can be applied to other language varieties to shed light on prosodic transfers.

6. Acknowledgements

7. References

[1]   Kloss, H., "'Abstand' languages and 'Ausbau' languages", Anthropological Linguistics, 9(7): 29–41, 1967.

[2]   Marcellesi, J.-B., "L'action thématique programmée : 'individuation sociolinguistique corse' et le corse polynomique", Études Corses, 28: 5–20, 1987.

[3]   Thiers, J., "Le français régional de Corse, une ressource ?", in M.-A. Maupertuis [Ed.], La Corse et le développement durable, pp. 99–105, Albiana, Aiacciu, 2010.

[4]   Dalbera-Stefanaggi, M. J., La langue corse, Presses Universitaires de France, Paris, 2002.

[5]   Filippi, P. M., Le français régional de Corse. Étude linguistique et sociolinguistique, PhD thesis, Università di Corsica, Corti, 1992.

[6]   Contini, M., Lai, J.-P., Romano, A., Roullet, S., de Castro Moutinho, L., Coimbra, R. L., Pereira Bendiha, U., Secca Ruivo, S. M., "Un projet d'Atlas Multimédia Prosodique de l'Espace Roman", 1st International Conference on Speech Prosody, Aix-en-Provence, pp. 227–230, 2002.

[7]   Durand, J. and Laks, B. "Relire les phonologues du français : Maurice Grammont et la loi des trois consonnes", Langue française, 126: 29–38, 2000.

[8]   Quenot, S., Structuration de l'École bilingue en Corse. Processus et stratégies scolaires d'intégration et de différenciation dans l'enseignement primaire, PhD thesis, Università di Corsica, Corti, 2010.

[9]   Bourdieu, P., Ce que parler veut dire. L'économie des échanges linguistiques, Fayard, Paris, 1982.

[10] Munro, M. J., "Nonsegmental factors in foreign accent: Ratings of filtered speech", Studies in Second Language Acquisition, 17: 17–34, 1995.

[11] Magen, H. S., "The perception of foreign-accented speech", Journal of Phonetics, 26: 381–400, 1998.

[12] Jilka, M., The contribution of intonation to the perception of foreign accent, PhD thesis, Universität Stuttgart, Stuttgart, 2000.

[13] Holm, S., Intonational and durational contributions to the perception of foreign-accented Norwegian: An experimental phonetic investigation, PhD thesis, Norwegian University of Science and Technology, Trondheim, 2008.

[14] Lai, J.-P., "Aires dialectales et intonation", Études Corses, 59: 95–110, 2005.

[15] van Bezooijen, R. and Gooskens, C., "Identification of language varieties. Contribution of different linguistic levels", Journal of Language and Social Psychology, 18(1): 31–48, 1999.

[16] van Leyden, K. and van Heuven, V. J., "On the prosody of Orkney and Shetland dialects", Phonetica, 63, 149–174, 2006.

[17] Boula de Mareüil, P. and Vieru-Dimulescu, B., "The contribution of prosody to the perception of foreign accent", Phonetica, 63: 247–267, 2006.

[18] Boula de Mareüil, P., Brahimi, B., Gendrot, C., "Role of segmental and suprasegmental cues in the perception of Magrebian-accented French", 8th International Conference on Spoken Language Processing, Jeju, pp. 341–344,.2004.

[19] Kaglik, A. and Boula de Mareüil, P., "Polish-accented French prosody in perception and production: transfer or universal acquisition process?", 5th International Conference on Speech Prosody, Chicago, pp. 1–4, 2010.

[20] Boula de Mareüil, P. and Lehka-Lemarchand, I., "Can a prosodic pattern induce/reduce the perception of a lower-class suburban accent in French?", 17th International Congress of Phonetic Sciences, Hong Kong, pp. 348–351, 2011.

[21] Ramus, F. and Mehler, J., "Language identification with supra-segmental cues: A study based on speech resynthesis", Journal of the Acoustical Society of America, 105(1): 512–521, 1999.

[22] D'Imperio, M. "Tonal alignment, scaling and slope in Italian question and statement tunes", 2nd Interspeech Event, Aalborg, pp. 99–102, 2001.

[23] Grice, M., D'Imperio, M., Savino, M., Avesani, C., "Strategies for intonation labelling across varieties of Italian", in Jun, S.-A. [Ed.], Prosodic typology: the phonology of intonation and phrasing, pp. 55–83, Oxford University Press, Oxford, 2005.