Rhythmic differences within Romance: identifying French, Spanish, European and Brazilian Portuguese

Andreas Dufter* and **Uli Reich****

* Universität München, Institut für Romanische Philologie
dufter@romanistik.uni-muenchen.de

** Universität zu Köln, Romanisches Seminar
uli.reich@uni-koeln.de

Abstract

Tradition has it that the Romance languages all share the property of being syllable-timed. More recent work, however, has challenged this view, emphasizing differences between languages and even between national varieties of a language like Portuguese.

To test whether French, Spanish, European and Brazilian Portuguese have their own prosodic signature, we conducted a perception experiment. We used low-pass filtered samples of carefully read, fast read, and elicited spontaneous speech as stimuli. These were presented to subjects familiar with at least one Romance idiom and who were asked to identify the language. We found that recognition was best for French, irrespective of speaking style, and second best for Spanish. Brazilian Portuguese was recognized at above pure chance, unlike European Portuguese. In general, identification was easier for spontaneous than for read speech. However, even for stimuli of the same language and same style, we found differences in recognizability depending on speaker-specific variables.

1 Introduction

Discussions of stress- vs. syllable-timing as mutually exclusive types of language rhythm tend to emphasize the difference between Germanic and Romance idioms, most often characterizing the latter as being all syllable-timed and, therefore, rhythmically alike. For French, however, the marked lengthening of syllable rhymes at the end of higher prosodic units is sometimes analyzed as a systematic demarcative strategy and is no longer dismissed as an occasional phonetic imperfection contrasting with a regular syllable-timed rhythmic ideal ([4], [9]). Moreover, some researchers have claimed that European Portuguese (henceforth EP) is closer to the stress-timing than to the syllable-timing type, due to the marked dominance of prominences, which licenses complex syllable structures in stressed syllables and various reductive processes elsewhere ([5]). In this manner, EP contrasts with Brazilian Portuguese (BP) as well as with European and American Spanish, idioms that share a more regular, alternating rhythm that causes the sequence of syllables to appear more equal and increases the salience and importance of syllabic prosody vis-à-vis phonological and morphophonological regularities ([12]).

However, for all languages characterized as syllable-timed, there has been little evidence for any preference toward isochronic perception of syllables, to say the least ([2]). Newer acoustic studies have sought alternative phonetic correlates, such as proportions and standard deviations of vocalic and consonantal interval durations ([7], [10]) or the distribution of sonority in the speech signal ([6]). Yet, some recent work advocates perceptual notions of rhythmic types, thereby making language discrimination and identification tasks central to phonetic research on language rhythm (e.g. [11]). In addition, experimental studies on systematic aspects of prosodic variation within languages have revealed strong effects resulting from speaking style ([8]) and, at least for some subjects, speaker-specific variables ([3], [7]).

Taking both speaker identity and speaking style into account, we tested to what extent listeners familiar with Romance languages, relying only on prosody, are able to identify speech as French, Spanish, EP or BP. Based on our own phonological work ([4], [12]), we hypothesized French to be more easily discernible than the rest due to final lengthening. Among the three Ibero-Romance varieties we considered, we expected best identification rates for EP, and massive confusion of Spanish and BP due to their similar accentual profile.

2 Method

2.1 Preparation of the stimuli

Samples of speech in three different styles were recorded from two native speakers of Metropolitan

ISBN 1-876346-48-5

French, three speakers of Standard Spanish from Bogotá, two speakers of EP and two speakers of BP (one from São Paulo, one from Rio). We asked our informants to read a short passage taken from Saint-Exupéry's "The Little Prince" – which we chose because of its relatively simple syntax – first in a slow, careful style, then in a fluent and rather quick style. Finally, we taped several minutes of free spontaneous speech involving topics they felt at ease speaking about. The recordings were done on minidisc and then stored on PC using PRAAT (at a sampling rate of 22 050 Hz). From each sample, we extracted twenty seconds of speech containing only declarative sentences and few hesitation phenomena. These signals were then low-pass filtered at 400 Hz using the PRAAT Hann band filtering option.

2.2 The perception experiments

We conducted two series of experiments, one at the University of Munich (with 19 subjects), the other at the University of Cologne (with 22 subjects). All participants reported normal hearing capacity. After a brief explanation of the task, 21 stimuli were presented to our subjects who were all advanced students or instructors in the departments for Romance languages. For each of our three speaking styles, we chose seven stimuli in order to avoid any assumption of equal distribution of our varieties. Participants had to fill out a questionnaire and try to identify, for each stimulus, the language (variety) as French, Spanish, EP or BP. Moreover, we asked them about their knowledge of the Romance varieties involved. They ranked their active competence from null to (near-)native and their passive exposure from none to daily.

In Munich, all participants were exposed to the stimuli simultaneously and could listen to the stimuli through headphones. In Cologne, however, each participant took part in the experiment on an individual basis and listened to the stimuli, which were played back through loudspeakers. On average, the interview sessions had a total duration of about 15 minutes each.

To be able to control potential learning effects during the course of the experiment, the session in Munich was organized differently from the one in Cologne. In neither case did we find an improvement in correct identification rates at later stages of the experiments. Furthermore, subjects reported after the experiment that they had not been able to develop an identification strategy.

3 Results

On the whole, our recognition results showed a significantly better than 25 % chance distribution: we counted 317 correct responses out of a total of 714 answers given, which amounts to 44 %. Even if we take into account not only the 397 incorrect, but also the 147 answers omitted, the proportion of correct responses still reached 37 %. Table 1 shows, however, that there are clear differences among speaking styles: Correct identification was easier with spontaneous speech than with read speech, carefully read speech was easier than quick reading.

	Σ *correct*	Σ *erroneous*	% *correct*
careful reading	122	149	45
quick reading	92	154	37
spontaneous	113	94	55

Table 1: Identification by speaking style

Next, we looked at the results for our four Romance varieties separately. French (Fr) attained the best rate of recognition, Spanish (Sp) and BP were reliably above the level of chance. By contrast, at least in our data, EP was not significantly recognizable. The respective figures are given in table 2:

	Σ *correct*	Σ *erroneous*	% *correct*
Fr	101	53	66
Sp	94	113	45
EP	40	110	27
BP	82	121	40

Table 2: Identification by language/variety

Taking into account both speaking style and language variety, we see that the respective counts do not always reflect the probabilitites that were to be expected. While in French identification performance improved strongly in spontaneous speech, this was not the case for our Ibero-Romance varieties, as is revealed in table 3:

	Σ *correct*	Σ *erroneous*	% *correct*
Fr, careful reading	43	34	56
Sp, careful reading	35	42	45
EP, careful reading	15	26	37
BP, careful reading	29	47	38
Fr, quick reading	19	13	59
Sp, quick reading	33	37	47
EP, quick reading	16	61	21
BP, quick reading	24	43	36
Fr, spontaneous	39	6	87
Sp, spontaneous	26	34	43
EP, spontaneous	9	23	28
BP, spontaneous	29	31	48

Table 3: Identification by style and variety

If we look at specific stimuli and take into account inter-speaker variation, we discover some noticeable correlations between speaker identity and identification performance: Stimulus 7 (carefully read Spanish)

ISBN 1-876346-48-5

and 14 (Spanish read quickly) were the overall winners, as can be seen from table 4, whereas stimulus 9 (Spanish read quickly) and 18 (spontaneous Spanish) had the worst rates of correct identification (figures of correct identification are given in boldface). Interestingly, 7 and 14 both stem from our first Colombian informant, 9 and 18 both from our second one. Another case to be studied individually is one of our EP stimuli (number 8) which was misidentified as Spanish by 62 % of our subjects (see again table 4). Even in our "worst cases" 8, 9 and 18, however, most subjects misidentified Spanish as EP or BP and vice-versa, but not as French. We will take up these issues in the discussion.

	Fr	*Sp*	*EP*	*BP*
stimulus 7	0	**29**	7	4
stimulus 8	1	24	**6**	8
stimulus 9	2	**4**	17	11
stimulus 14	1	**29**	3	3
stimulus 18	4	**5**	15	6

Table 4: Best and worst identification for single stimuli

4 Discussion

Generally, the number of correct language identification indicates that this task was neither too difficult nor too easy, since our results only rarely come close to chance distribution and never to 100 % recognition. Prosody seems to vary enough among the Romance varieties discussed here to allow for some degree of identification. As can be seen from table 1, on the whole, prosodic differences turn out to be clearer in spontaneous speech than in read speech. Thus, our experiment confirms that reading levels out differences between language-specific prosodic characteristics ([1]). We had expected the fast reading condition to be the most difficult for our identification task, since, cross-linguistically, higher tempo should reduce the salience of lower prosodic units such as the syllable and the phonological word in favour of larger accentual phrases ('stress-timing'). Moreover, quicker reading should weaken phrase-final lengthening which we expected to be important for the discrimination of French. However, our data indicate that careful reading has nearly the same effect of levelling out language-specific differences, since the advantage of the careful reading condition is almost exclusively due to a single EP stimulus (number 8, represented in table 4), which completely fell out of the general picture. The reason for this higher uniformity of slow reading speech might be that the opposite trend, namely the higher salience of lower units in slower read speech, is also a cross-linguistic trend ('syllable-timing').

When we discuss stimulus 8 separately, the dependency of the prosodic profile on discourse-type and individual performance becomes visible: it was realized by a Portuguese language teacher, who tends to produce overly clear speech, pronouncing vowels which, in less monitored EP speech, are likely to be omitted. This might have led to the clear misidentification of a EP stimulus as Spanish (24 out 39, i. e. 62 %), a language with clearly pronounced vowels and simple syllable structures.

Across all speaking styles, identification of French was best, followed by Spanish and BP. Does this asymmetry reflect different degrees of prosodic markedness or rather that the subjects possessed different degrees of familiarity with these varieties? We tried to reach a tentative answer by selecting subsets of our informants who reported excellent or native competence and/or high familiarity. Among the 41 participants, we had 20 (near-)native listeners of French, 20 of Spanish, and only 4 of (European or Brazilian) Portuguese. Table 5 shows percentages of correct identification of (near-)native languages by the respective subsets of our listeners:

	Σ *correct*	Σ *erroneous*	% *correct*
Fr	40	43	48
Sp	44	70	39
EP/BP	13	11	54

Table 5: Identification of (near-)native languages

Comparing these results to the overall recognition proportions in table 2 above, we find that recognition of Portuguese by Lusophone subjects is much better than average, whereas French and Spanish recognition rates fall even further. This looks astonishing at first, yet the reason for this could well be the skewed distribution of linguistic competence among our subjects: While our Lusophone subjects mostly know some Spanish and French as well, the French and Spanish (near-)natives in both groups of subjects tend to have little or no familiarity with Portuguese, neither with its European nor with its Brazilian variety.

Among the erroneous language identifications, we found that (near-)native listeners of French and Spanish very rarely confused these two languages: Only 11 identified French stimuli as Spanish, and only 6 Spanish stimuli as French. The Portuguese results were less clear, however. In a similar perception experiment ([5]), it was found that Portuguese subjects, while not being able to choose correctly between EP and BP on the basis of low-pass filtered and F0-flattened stimuli, could do so with filtered stimuli where F0 was kept intact. We could not replicate these findings, since even with F0 kept intact, our subjects, on the whole, discriminated poorly between EP and BP.

The extremely poor identification of the Spanish stimuli 9 and 18, both from the same speaker, might be explained by reference to F0, which is extremely

ISBN 1-876346-48-5

high, thus giving a particularly strange impression. Since many of our subjects are least familiar with Portuguese, they overwhelmingly misidentified the stimuli as EP or BP, which were the strangest options available. It is possible that this 'strangeness effect' could have been eliminated by offering an exotic language as a further possibility.

Contrary to our hypotheses, Spanish and BP were not especially liable to be confused, although they show similar degrees of syllable structure complexity and possess similar stress patterns. We are planning to reconsider this issue by carrying out language discrimination tasks using Ibero-Romance language pairs.

We deliberately refrained from flattening F0 contours, as carried out in several studies (e. g. [5], [11]), which, of course, leaves open the possibility of intonational cues for language identification besides rhythmic ones. However, the more complete prosodic picture retains more language-like properties and should therefore be more likely be perceived as language. This, in our view, provides a considerable advantage, since we had to cope with our subjects' tendency to attribute 'strange sound files' to languages least familiar to them (such as Portuguese in our experiment).

5 Conclusion

Our investigation differs from similar language identification studies by making the subjects choose between more than two idioms. This task is, of course, more demanding than discrimination of language pairs, yet some interesting findings emerge only in an experimental design providing more than one false answer, e.g. looking at cases of consistent misidentification. However, with more languages to identify, the subjects' individual degrees of familiarity with those languages have to be considered more carefully.

In general, correct language identification was clearly at above chance probability in our experiments, especially in the spontaneous speech samples. Reading, however, reduced identifiability at both quick and slow speech rates. Therefore, the influence of speaking style on recognition cannot exclusively be due to systematic differences in speech rate. Prosodic characteristics are not only dependent on the language, but also on the type of discourse. Besides temporal cues, the relative salience of different prosodic units and the distribution of accentual prominences seem to be crucial for language(-type-)specific rhythms. Since syllables and accentual domains are more properly regarded as phonological notions, we conclude by suggesting that perception experiments like ours can only give an indication of language-specific prosodic differences, whereas the rhythmic organization of a language has to be described within the domain of phonology.

Acknowledgments

We would like to thank our native speakers and our participants in the perception experiments. Special thanks go to Damon Davison for checking our English.

References

[1] Marie Bernadete Marques Abaurre. "The rhythms of speech and writing", In Clotilde Pontecorvo, Margherita Orsolini, Barbara Burge, Lauren B. Resnick (eds.), Children's early text construction. Mahwah, NJ: Erlbaum, 1996, pp. 47–66.

[2] Peter Auer. "Silben- und akzentzählende Sprachen als phonologische Typen", In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, Wolfgang Raible (eds.), Sprachtypologie und sprachliche Universalien/ Linguistic typology and universals. Berlin/New York: de Gruyter, 2001, vol. 2, pp. 1391–1399.

[3] Z. S. Bond, Verna Stockmal, D. Muljani. "Learning to identify a foreign language", In Language Sciences 20, 1998, pp. 353–367.

[4] Andreas Dufter. Typen sprachrhythmischer Konturbildung, Tübingen: Niemeyer, in press.

[5] Sónia Frota, Marina Vigário. "On the correlates of rhythmic distinctions: The European / Brazilian Portuguese case", In Probus 13, 2001, pp. 247–275.

[6] Antonio Galves, Jesus Garcia, Denise Duarte, Charlotte Galves. "Sonority as a basis for rhythmic class discrimination", In Proc. Speech Prosody 2002, Aix-en-Provence.

[7] Esther Grabe. "Variation adds to prosodic typology", In Proc. Speech Prosody 2002, Aix-en-Provence.

[8] Isabelle Guaïtella. "Rhythm in speech: what rhythmic organizations reveal about cognitive processes in spontaneous speech production vs. reading aloud", In Journal of Pragmatics 31, 1999, pp. 509–523.

[9] Anne Lacheret-Dujour, Frédéric Beaugendre. La prosodie du français, Paris: CNRS éditions, 1999.

[10] Franck Ramus, Marina Nespor, Jacques Mehler. "Correlates of linguistic rhythm in the speech signal", In Cognition 73, 1999, pp. 265–292.

[11] Franck Ramus, Jacques Mehler. "Language identification with suprasegmental cues: A study based on speech resynthesis", In J. Acoust. Soc. Am. 105, 1999, pp. 512–521.

[12] Uli Reich. Freie Pronomina, Verbalklitika und Nullobjekte im Spielraum diskursiver Variation des Portugiesischen in São Paulo, Tübingen: Narr, 2002.