Language Discrimination and Rhythm Classes: Evidence from Portuguese

Sónia Frota, Marina Vigário** & Fernando Martins**

*DLGR & Lab. de Fonética, University of Lisbon, Portugal
**ILCH, University of Minho, Portugal
{sonia.frota; marina.vigario}@mail.telepac.pt; fmartins@netcabo.pt

Abstract

This paper reports on two discrimination experiments involving Portuguese data. The findings support the prosodic contrast between EP and BP and stress the role played by intonation in this contrast. They also show that the perceptual distance between EP and BP is smaller than that between EP and Dutch. The robust distinction between EP and Dutch provides evidence for different perceptual weights of acoustic parameters, and points to the syllable-timed nature of EP.

1. Introduction

Until recently, the literature on the rhythm of Portuguese classified European Portuguese (EP) as stress-timed and Brazilian Portuguese (BP) as having mixed patterns of the syllable and stress-timed type (e.g. [12]). In either case, no clear support for this classification, experimental or other, was given. Largely following the approach to rhythm put forward in [14], [5] presents an account of Portuguese rhythm based on acoustic measures of consonantal and vocalic intervals and explores the relation between these measures and the phonological properties specific to each variety. It is shown that the acoustic parameters %V (proportion of vocalic intervals) and ΔC (variability of consonantal intervals, expressed by a standard deviation measure) successfully capture the rhythmic distinction between EP and BP: EP is clearly 'more' stress-timed, whereas BP is clearly 'more' syllable-timed. In addition, a comparison with the results for 8 other languages reported in [14] allows EP and BP to be related with the traditional rhythm typology as follows: EP has **both** stress-timing (ΔC) and syllable-timing (%V) properties, whereas BP shows **both** syllable- (ΔC) and mora-timing (%V) properties. This is shown in Fig. 1.

The findings in [5] raise two interesting questions. It is known that infants use rhythm to discriminate between languages (e.g. [10]) and it has been proposed that rhythm may help in the acquisition of certain phonological properties, such as syllable structure ([14]). In this view, rhythm type may be used as a cue to other properties of the language. In languages like EP and BP, the speech signal provides the native speaker with mixed cues relative to the rhythm of the language. If both stress- and syllable-timing cues are contained in the signal, like in the EP case, how can rhythm provide that help?

A different but related issue is the perceptual weighting of the acoustic parameters involved. From the data reported in [14], it was not clear how the relative importance of vocalic space and consonantal interval variability to the perception of rhythm could be tested, as the two parameters always converged on the same classification (Fig.1). That is, however, not the case in EP and BP. Thus, the relevance of each of the two parameters may be assessed on the basis of the Portuguese data.

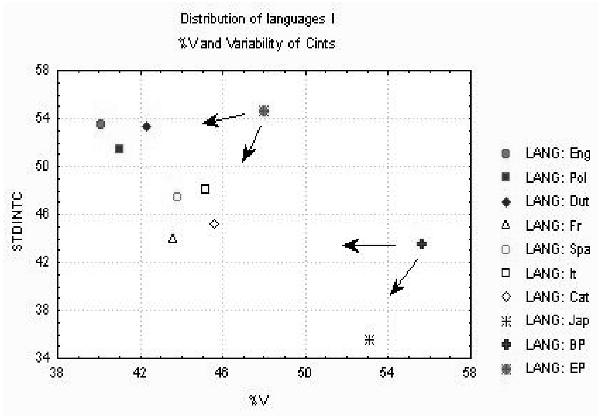

Figure 1: Distribution of languages over the %V, ΔC plane (from [5]; data for the 8 other languages from [14]).

This paper reports on two of a series of discrimination experiments designed to address these questions. It has two conjoint goals: (i) to test whether EP and BP are discriminated on the basis of prosodic information, in several experimental conditions; (ii) to investigate the perceptual weighting of %V and ΔC, and the relative perceptual distance between EP and other languages.

Besides %V and ΔC, the relevance of intonation to language discrimination was also considered. Although in a previous study on the role of various prosodic cues intonation was found to be neither necessary nor sufficient for language discriminability ([13]), other work on the role of intonation has reached a different conclusion (e.g. [15]). It is known that EP and BP have remarkably different intonational properties ([4]), as the contours in Fig. 2 illustrate. The possibility that intonation may contribute to discriminate EP from BP, or each of the two from other languages, should not be dismissed without investigation.

2. Experiment 1: EP / BP

2.1. Methods

Sentences were selected from a comparative EP/BP corpus composed of the Portuguese translation of the multi-language corpus used in [14] (see [5]). All sentences are short, simple declaratives read as news sentences in a soundproof booth by female native speakers of EP and BP. The sentences selected contain 15 to 19 syllables each and are representative of the difference between EP and BP concerning the acoustic parameters %V and ΔC.

The source sentences were low-pass filtered using a filter with a 400-Hz frequency cutoff.

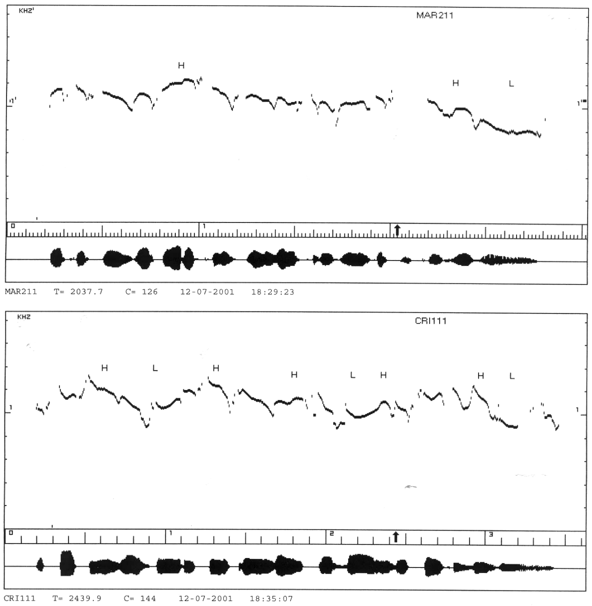

Figure 2: F0 contours of the sentence *A catalogadora compreendeu o trabalho da pesquisadora* ‘The catalogue-maker understood the work of the researcher’, uttered by an EP (upper) and a BP speaker (lower panel).

Two versions of the filtered sentences were created: one corresponds to the output of the filtering transformation; in the other the F0 contour was made flat, taking the mean F0 value of the original contour. These transformations were made using PRAAT. For each of the two conditions ‘+F0’ and ‘-F0’, 16 pairs of sentences were built, 6 of each containing stimuli from the same language (either EP or BP) and 10 combining stimuli from EP and BP. The stimuli in each pair were always uttered by two different speakers.

The AX discrimination task was presented to 29 subjects. The instructions indicated that they would be listening to acoustically modified sentences of two exotic languages, Tigre and Hua. The task was preceded by a training phase. The training consisted of 4 steps: a set of sentences from Tigre (which was EP), a set of sentences from Hua (BP), a set of AX pairs where X is the same as A followed by a set of AX pairs where X is different, and finally a set of AX pairs of both types presented in random order. The subjects had feedback during the training and could only do the training once. The sentences in the training set were not used in the task. The training and the task lasted about 15 minutes for each of the two conditions tested.

The subjects were university students, with a mean age of 26 years. They were all EP native speakers. They were tested simultaneously in a language lab provided with individual listening booths. When asked, after the experiment, to rate the degree of difficulty of the task, most subjects considered the task with ‘+F0’ stimuli difficult and the task with ‘-F0’ stimuli very difficult. In no case were any of the real languages at stake disclosed.

2.2. Results

It is known that not all subjects are equally able to perform discrimination tasks with degraded stimuli, and not all sentences are equally representative of the rhythmic properties of the language (e.g. [13]). We tried to obviate the second problem by selecting source sentences that are representative of the language with regard to the acoustic parameters related to rhythm. As to the first problem, the following rejection criterion was set: subjects that could not identify the pairs with stimuli drawn from the same language as the SAME above chance level were excluded from the analysis. The results reported below include data from 21 subjects for the ‘+F0’ condition, and from 12 subjects for the ‘-F0’ condition.

	‘+F0’	‘-F0’
Percent scores	62.38	46.67
U and Z values	U=94.5; Z=3.64	U=66; Z=-.43
p-level	.0003	.67
d’-scores	1.88	1.28

Table 1: Mean percent scores, statistical results of the scores tested against chance level (50%), and d’-scores.

Table 1 presents the discrimination results obtained averaged across subjects. Percent scores correspond to the total hit rate of all the DIFFERENT answers for the EP/BP pairs. A Mann Whitney U test was used to test the significance of the scores relative to chance level. True d’-scores were also computed: the DIFFERENT responses to the different pairs were used as hit rate and the DIFFERENT responses to the same pairs as false alarms ([8]). As can be seen, EP and BP are successfully discriminated only in the ‘+F0’ condition. With flat intonation, subjects’ responses are not different from chance. Intonation has thus significantly influenced discrimination.

2.3. Discussion

The results show that subjects can discriminate EP from BP on the basis of filtered sentences that have a %V and ΔC contrast, and in which the intonation pattern is preserved. However, when intonation is not a cue available to subjects, the perceptual distance between the two languages shortens (d’ is used as a measure of distance - see Table 1) and they are no longer reliably differentiated. Hence the results support the relevance of intonation to the discrimination between EP and BP, by assigning it the status of a necessary cue.

3. Experiment 2: Language categorization

3.1. Methods

A set of Dutch (DU) and Spanish (SP) sentences was selected from the multi-language corpus used in [14]. DU and SP stand as examples of stress- and syllable-timed languages, respectively. The EP and BP material was a subset of that used in Experiment 1. All the sentences had either 15 or 17 syllables, and the sentences within a pair were matched on their number of syllables. Construction of the materials was the same as for Experiment 1. Twenty pairs of sentences were built, 4 containing stimuli from the same language (either DU or SP), and 16 combining stimuli from different languages: 4 pairs of the DU/SP type, 4 pairs of the DU/EP type, 4 pairs of the DU/BP type, and 4 pairs of the EP/BP type.

The procedure followed was identical to that described in Experiment 1. In the training phase, the subjects listened to sentences from Tigre (which was Dutch) and Hua (which was Spanish). No sentences from EP or BP were included in the training. Consequently, the subjects not only had to generalize what they learned in the training to novel sentences of Dutch and Spanish, but also to decide whether EP and BP are of the Tigre or the Hua type. In other words, the task tests their ability to categorize DU, SP, EP and BP into two groups: the 'Tigre'-group and the 'Hua'-group. The rhythmic chart given in Fig. 1 above shows that BP has some acoustic properties of the Spanish-'Hua' type, whereas EP has properties both of this type and of the Dutch-'Tigre' type. If these properties are used by subjects, BP is predicted not to fall within the 'Tigre'-group. As to EP, it may either belong to the 'Tigre' or the 'Hua'-group depending on the acoustic parameter that is perceptually more salient. A third possibility would be an inconsistent classification of EP, what would mean that none of the parameters outweighs the other and thus EP does not fit into either of the groups. With regard to intonation, it is known that EP and Dutch may display similar patterns, namely the hat pattern contour ([9], [3]).

Thirty subjects participated in the experiment, with a mean age of 25 years. They rated the task as difficult or very difficult. In no case were any of the real languages at stake disclosed.

3.2. Results

The same rejection criteria described in Experiment 1 was applied. The results reported here include data from 7 subjects for the '+F0' condition, and data from 11 subjects for the '-F0' condition.

'+F0'	DU/SP	DU/EP	DU/BP	EP/BP
% score	67.86	64.29	60.71	17.86
U and Z	14;1.57	10.5;2.28	14;1.88	3.5;-3.00
p-level	.116	.022	.059	.003
d'-score	2.59	2.15	2.01	0.23
'-F0'	DU/SP	DU/EP	DU/BP	EP/BP
% score	52.27	59.09	52.27	18.18
U and Z	55;.49	44;1.32	55;.438	11;-3.69
p-level	.622	.188	.662	.0002
d'-score	1.94	2.39	2.09	0.29

Table 2: Mean percent scores, statistical results of the scores tested against chance level (50%), and d'-scores.

The results for the '+F0' condition are given in the first section of Table 2. It can be seen that all language pairs are discriminated (with scores of 60% or more) to the exception of EP/BP. The Mann Whitney U test results show that subject responses are not significantly different from chance level for DU/SP, and are near to significance for DU/BP. A detailed inspection of the data showed that the DU/SP non-significant result is mainly due to the behavior of one subject that shows the reverse pattern of that of the majority. EP is successfully discriminated from DU, whereas it is consistently not distinguished from BP.

Let us now consider the results obtained in the '-F0' condition. It is clear from the Mann Whitney U test results that the languages are generally less consistently discriminated in this condition than when intonation is preserved. The exclusion of the intonation cue, however, does not affect equally all language pairs. There is a substantial drop in discrimination for the DU/SP pairs, in particular, that is reflected both in the percent score and the d'-score. By contrast, the DU/EP pairs are much less affected. In fact, this is the only case that shows a clearly higher d'-score in the flat intonation condition than in the '+F0 condition'.

A perceptual map of the languages studied, using the d' values as a measure of distance, is displayed in Fig. 3 (see [7]). The position of SP in the perceptual space is not well specified, as no measure of its distance to either EP or BP is available. Nevertheless, it was included in the map as a point of comparison to the distances of EP and BP from DU.

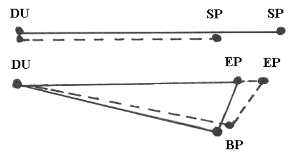

Figure 3: The four languages in perceptual space ('+F0': heavy line; '-F0': dashed line).

The perceptual map highlights the effect of intonation for DU/SP (an effect similar to that described in Experiment 1 for EP/BP): the perceptual distance between the two languages shortens without intonation. The map also shows an effect on the opposite direction for DU/EP: the perceptual distance between the two languages gets larger without intonation. The relative position of the other language pairs is not affected by the F0 condition.

3.3. Discussion

The results of Experiment 2 show that both BP and EP are discriminated from DU. This means that the subjects have differentiated them from the Dutch-'Tigre' type. This result is as expected for BP, due to its acoustic properties shown in Fig. 1. As to EP, this result constitutes a strong indication that the acoustic parameters at stake are not equally salient. EP and DU are distinguished by %V, but not by ΔC. The non-inclusion of EP in the Dutch-'Tigre' group can thus be seen as evidence for the perceptual salience of %V relative to ΔC.

The fact that the loss of intonation does not hinder discrimination in the DU/EP case, but promotes it instead, is consistent with the available descriptions of the neutral declarative intonation patterns of both languages. As DU and EP display similar patterns, this prosodic cue does not help in discriminating EP from DU. On the contrary, its absence enhances the prosodic distance between the two languages. The fact the intonation plays no role here further highlights the importance of the timing parameter expressed by %V.

Another result of Experiment 2 is the non-discrimination of EP and BP. Apparently, this result is at odds with that obtained in Experiment 1. The design of Experiment 2, however, provides a straightforward explanation for this apparent inconsistency. The task requires that the subjects categorize the stimuli into two groups. As both EP and BP are not included in the Dutch-'Tigre' group, only two possibilities are left: either they are both included in the other group, or there is some difference between the two languages that makes subjects unable to successfully fit them into the two groups at stake. The result obtained clearly shows that EP and BP are treated as the 'same' language. This is only consistent with their inclusion in the same group.

As to the DU/SP results, two comments are in point. It was seen the two languages are not always reliably discriminated, contrary to our initial expectations based on the clear stress-timed and syllable-timed status respectively of Dutch and Spanish in the traditional rhythm typology (e.g. [11], [14]). However, the results of simulated discrimination experiments with adults using %V as the relevant parameter presented in [14] show that DU/SP is the only deviant pair of the 26 pairs considered (EP and BP are not included in this sample): the score for DU/SP is lower than all the other between rhythmic class scores. Therefore, the result we obtained cannot be seen as peculiar. But more importantly, in the light of our results it appears that intonation has a role to play in DU/SP difference. We thus hypothesize that the inclusion of intonation as a relevant parameter would boost the discrimination score for this language pair.

4. General Discussion

In this study, we have presented data showing that European Portuguese adults are able to discriminate between filtered sentences drawn from different languages in certain conditions. As prosodic information is spared by the filtered nature of the stimuli, unlike detailed segmental information, we assume that discrimination was based on prosodic information such as the acoustic parameters %V and ΔC, which cue rhythmic distinctions, and the intonation pattern.

The average discrimination scores obtained in Experiments 1 and 2 in the '+F0' condition are within the range of those reported in other language discrimination studies, like [13] and [14] (between 60% and 70%). The overall drop in the scores in the '-F0' condition is also consistent with the increased difficulty imposed by highly impoverished stimuli. The performance of our subjects is thus in line with that of other subjects of similar tasks.

Experiment 1 demonstrates that EP and BP are discriminable only when the intonation pattern is preserved. This result is somewhat surprising in the face of the acoustic distance between the two languages in the rhythmic chart given in Fig. 1. However, it may well be the case that intonation is a very salient and thus necessary cue to the rhythmic difference between the two languages, at least for EP subjects. That intonation may be one of the important factors that lead to rhythmic distinctions has been suggested in [1], [2], among others. This result certainly calls for further research on the link between timing and intonation properties. The DU/SP results in Experiment 2 reinforce this need.

Experiment 2 shows that both EP and BP are discriminated from Dutch. The discrimination of EP from DU provides evidence for the perceptual weighting of %V and ΔC: the proportion of vocalic intervals is the most salient parameter. Experiment 2 also shows that, when contrasted to Dutch, EP and BP are no longer discriminated. This result suggests that the distance between EP and Dutch is bigger than the distance between EP and BP. The rhythmic chart in Fig. 1, which is acoustically based, does not mirror these relative distances. However, they are directly reflected by the d' values obtained, which we use as a measure of perceptual distance. It is thus conceivable that the smaller perceptual distance between EP and BP that is captured when just the two languages are compared fades away when confronted with the bigger perceptual distance that separates EP (and BP) from Dutch. These results are a strong indication that EP is not a stress-timed language, like Dutch, but a syllable-timed language instead.

Given the results discussed in the present work, it seems desirable that language discrimination studies are not only based on just two languages, but also include experiments that comprise a wider range of languages. In future work, we intend to follow this line of research, by assessing the location of Spanish in the perceptual space with more precision and by taking into account other languages. A further development is to control for a putative native language effect on the discrimination results (see [6]). This will be done by running similar experiments with non-native European Portuguese subjects. Among other aspects, the examination of the native language factor will allow us to determine whether the sensitivity to intonation reported above is a general property of the languages involved or a specific property of the European Portuguese native speaker's phonological system.

5. References

[1] Dauer, R., 1983. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11, 51-62.

[2] Dauer, R., 1987. Phonetic and Phonological Components of Language Rhythm. In *Proceedings of the XIth International Congress of Phonetic Sciences*, 268-274.

[3] Frota, S., 2000. *Prosody and Focus in European Portuguese*. New York: Garland Publishing.

[4] Frota, S.; Vigário, M., 2000. Aspectos de prosódia comparada: ritmo e entoação no PE e no PB. In *Actas do XV Encontro da APL*. Braga: APL, 533-555.

[5] Frota, S.; Vigário, M., 2001. On the correlates of rhythmic distinctions: The European / Brazilian Portuguese case. *Probus* 13, 247-275.

[6] Hume, E.; Johnson, K., 2001. A Model of the Interplay of Speech Perception and Phonology. In *The Role of Speech Perception in Phonology*, E. Hume & K. Johnson (eds.). New York: Academic Press, 3-26.

[7] Johnson, K., in preparation. *Acoustic and Auditory Phonetics*. 2nd ed.

[8] Kaplan, H.L.; Macmillan, N.A.; Creelman, C.D., 1978. Methods & Designs. Tables of d' for variable-standard discrimination paradigms. *Behavior Research Methods & Instrumentation* 10(6), 796-813.

[9] Ladd, D.R., 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.

[10] Nazzi, T.; Bertoncini, J.; Mehler, J., 1998. Language discrimination by newborns: towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance* 24(3), 756-766.

[11] Nespor, M., 1990. On the rhythm parameter in phonology. In *Logical Issues in Language Acquisition*, I. Roca (ed.). Dordrecht: Foris, 157-175.

[12] Parkinson, S., 1988. Portuguese. In *The Romance Languages*, M. Harris & N. Vincent (eds.). London: Routledge, 131-169.

[13] Ramus, F.; Mehler, J., 1999. Language identification with suprasegmental cues: A study based on speech resynthesis. *JASA* 105, 512-521.

[14] Ramus, F.; Nespor, M.; Mehler, J., 1999. Correlates of linguistic rhythm in speech. *Cognition* 73, 265-292.

[15] Willems, N., 1982. *English intonation from a Dutch point of view*. Dordrecht: Foris.

Sound samples can be found at <http://www.fl.ul.pt/pessoais/sfrota/soundfiles.htm>. Thanks are due to M. Beckman, K. Johnson, B. Ladd and F. Ramus. This work was supported by the FCT / FEDER project POCTI / 33277 / LIN / 2000.