CATEGORICAL PERCEPTION OF *PORTEÑO* NUCLEAR ACCENTS

Ingo Feldhausen[a,b]*, Andrea Pešková*[a]*, Elena Kireva*[b] *& Christoph Gabriel*[a,b]

[a]Collaborative Research Center 538 'Multilingualism' Hamburg, Germany;
[b]University of Hamburg, Germany
ingo.feldhausen@gmx.de; christoph.gabriel@uni-hamburg.de

ABSTRACT

By applying the categorical perception (CP) paradigm complemented by measurements of reaction time (RT), we provide evidence for a categorical scaling contrast in the nuclear position of utterances with narrowly focused constituents in Argentinean *Porteño*-Spanish. While the tritonal accent L+H*+L is perceived as signaling contrast, L* is interpreted as its neutral counterpart.

Keywords: Argentinean Spanish, tritonal pitch accent, contrast, focus, categorical perception

1. INTRODUCTION

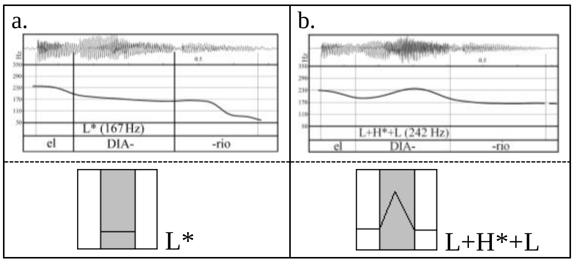
The presence of the tritonal pitch accent L+H*+L in the tonal inventory of *Porteño*-Spanish represents an important characteristic of this Argentinean variety [4]. L+H*+L is characterized by a rise and fall within the limits of the stressed syllable (Fig. 1b); it represents the typical nuclear accent in declaratives with a contrastively focused constituent in clause-final position (ex. 1b/2; Fig. 1b). The nuclear accent L* is typical of statements with clause-final neutral focus (1a/2; Fig. 1a).

(1) a. Neutral context
¿Qué le dio María a su hermano?
'What did Mary give her brother?'

b. Contrastive context
María le dio una revista a su hermano, ¿no?
'Mary gave a magazine to her brother, right?'

(2) Response to 1a and b (target phrase marked by […])
María le dio a su hermano [*el DIArio*].
'Mary gave a newspaper to her brother.'

Figure 1: Base stimuli and schematized pitch accents (1a: neutral context; 1b: contrastive context).



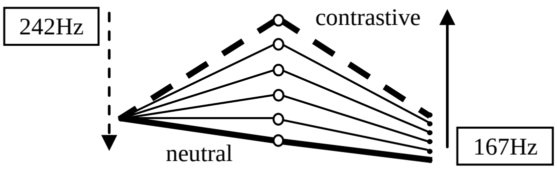
Until now, the formal and functional distinction between L* and L+H*+L has been based on production data. In the following, we report on a perception experiment conducted in order to test whether the difference in pitch height between the two accents is indeed categorical. Using the CP paradigm, which consists of an identification task and a discrimination task [9], we follow the still young tradition of applying this method to prosodic research [6], [8], [12]-[14]. In the identification task, the listeners' task was to classify stimuli taken from a continuum; in the discrimination task, they were asked to determine whether pairs of stimuli were identical or different. We take the following assumption as a point of departure: Provided that L+H*+L nuclear accents are typical of emphatic/contrastive statements in *Porteño*-Spanish, listeners should make strict linguistic use of F0 scaling differences in the perception of statements involving contrastiveness as opposed to statements with neutral narrow focus. Following [2], we also measure the mean reaction time (RT) necessary for identification. We hypothesize that the mean RT is shorter for within-category stimuli than for across-category stimuli, since the former are expected to require less cognitive effort (in the form of decision making) than the latter [2].

2. METHODOLOGY

We conducted two CP experiments combined with RT measurements. To set up the experiment, we used the DMDX software, a display system used to measure RTs to visual and auditory stimuli [3]. RTs were measured from stimulus onset until the participant's response. Both experiments are based on target sentence (2), produced by a female native speaker of *Porteño*-Spanish as a reply to either (1a) or (1b). In the first experiment, listeners were confronted with the whole sentence (2) in its original and an F0 manipulated form (see below); in the second experiment, they only heard the object *el diario* 'the newspaper'. The reason for this is the fact that contrastive/emphatic statements are characterized by prosodic cues other than F0, such as higher degrees of intensity and sonority than in neutral ones [7]. These cues might interact with the F0 manipulations performed on the nuclear accent and confuse the listeners. For example, when a stimulus originally produced in the contrastive context (1b/2) receives its nuclear accent manipulated from L+H*+L towards L*, it signals emphasis/contrast in the prenuclear utterance section (through its high intensity and sonority), while the F0 shape of the nuclear accent conveys a neutral interpretation. This effect is minimized when the listener is only confronted with the object.

The original peak was at 242Hz in the contrastive context and at 167Hz in the neutral one (Fig. 1). The stimuli were stylized (2 semitones) using *Praat* [1]; the F0 scaling was manipulated in increments of 15Hz. Five tokens were created by lowering the peak from the original contrastive stimulus; five more tokens were created by shifting the peak upwards from the neutral base stimulus (empty circles in Fig. 2), resulting in a total of 12 stimuli (2 x 5 manipulations, 2 original stimuli). Differences in F0 at the right edges of the accented syllable were compensated for by including the right elbow in the manipulation (filled circles).

Figure 2: Schematic representation of the manipulations performed on original stimuli (bold solid line: neutral context; bold dashed line: contrastive context).



In the identification task, the 12 stimuli were repeated five times and presented in blocks of 12 in a block-internally randomized order. The listeners were asked to classify each stimulus as appropriate for either the neutral or the contrastive context by pressing the corresponding button of the keyboard. The pairs of stimuli from the identification task were also used for the discrimination task. Ten pairs were created in AB order, with changes in an upwards direction (steps 1+2, 2+3 etc.); ten more pairs contained the same stimuli in the reverse order (BA, i.e. downwards change; steps 2+1, 3+2 etc.). Twelve control pairs with identical items were added (AA, steps 1+1, 2+2 etc.). For all of the 32 pairs obtained, half of the items stemmed from the contrastive stimulus and the others from the neutral one. The stimulus pairs were repeated three times and presented in blocks of 16 items in a block-internally randomized order. The listeners were asked to judge whether the pair they heard consisted of either identical or different stimuli. Both tasks started with an introductory explanation and a practice session.

Eight speakers of *Porteño*-Spanish living in Hamburg, Germany, participated in the experiment (7 males, 1 female, ages 27-33 [one male 69]; no hearing impairments). Two of them, however, were heritage speakers (HS) and behaved differently than the others. Their data are not included in the results for reasons discussed in section 4.

The two experiments were conducted in a variable order. Half of the listeners were confronted with the order sentence-object (i.e. 1a. ident.-object, 1b. ident.-sent.; 2a. discr.-object, 2b. discr.-sent.), the others with the reverse order (i.e. 1a. ident.-sent., 1b. ident.-object; 2a. discr.-sent., 2b. discr.-object).

3. RESULTS

3.1. Identification Task

Fig. 3 shows the percentages of ‘contrastive’ responses for both experiments. Low scaling of the nuclear accent leads to low percentages of ‘contrastive’ judgments, while a high scaling leads to high percentages of ‘contrastive’ responses. The S-shape of the function is clearer in the lower panel than in the upper one and is even more obvious for the continuum created on the basis of the contrastive stimulus (dashed line). The solid line in Fig. 3a rises slower than that in Fig. 3b. This means that the ‘contrastive’ interpretation of the sentential stimuli starts later than that of the object stimuli. In addition, the percentages of ‘contrastive’ responses are higher for the object than for the sentential stimuli (212, 227, 242Hz). The same holds for the manipulations stemming from the contrastive stimulus. The sentential stimuli were judged more often as being contrastive at the lower peak values than the object stimuli. Fig. 3 also plots the standard error (SE) of the identification rate, defined as the number of ‘contrastive’ responses for each continuum. Stimuli 167 and 242Hz had low SEs (average values: ±5.87 in Fig. 3a, ±3.53 in 3b); the stimuli in the crossover area (197, 212Hz) generally exhibited higher SEs (±14.17 in Fig. 3a, ±11.67 in 3b). Only the 197Hz stimulus (Fig. 3a) deviated from the pattern, with a SE value of ±4.2.

The mean RTs are given in Fig. 4. The dashed lines have (almost) the ideal course for the task. Classification of the stimuli took more time between categories (with a peak at 197Hz; very clear in Fig. 4b) than within them. The solid line is also nearly optimal in Fig. 4a (with a peak at 212Hz), but shows a sudden rise at 242Hz. In Fig. 4b, this shows that listeners can rapidly classify the 167Hz stimuli, in sharp contrast to the other stimuli.

Figure 3: Percentages of 'contrastive' responses on the continuum of test items created from the neutral (solid line) and contrastive (dashed line) stimuli (upper panel/Fig. 3a: sentence; lower panel/Fig. 3b: object).

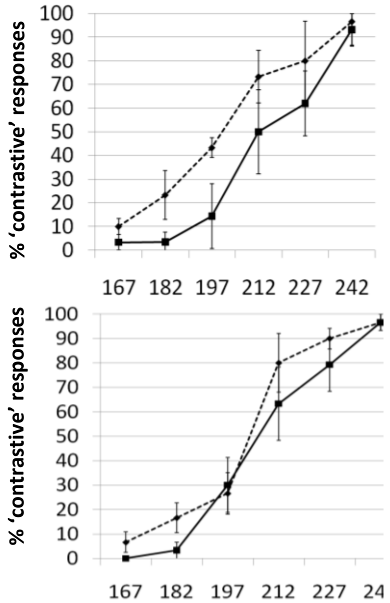
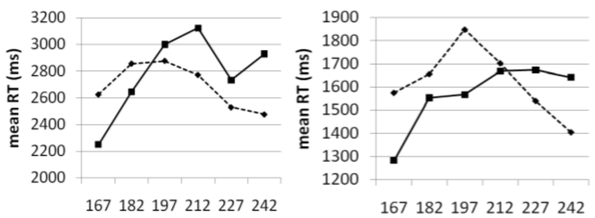


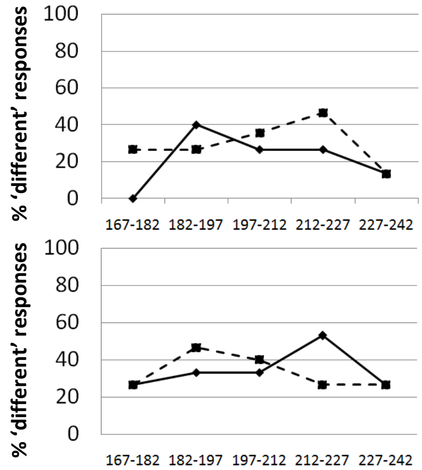
Figure 4: Mean RTs in ms on the continuum of test items created from the neutral (solid line) and contrastive (dashed line) base stimuli (left panel/Fig. 4a: sentence; right panel/Fig. 4b: object).



3.2. Discrimination Task

Fig. 5 illustrates the results for the AB order. All conditions show a clear discrimination peak. However, it is never located directly at 197-212Hz, instead either immediately to its left (182-197Hz) or right (212-227Hz), with a discrimination rate of about 50%.

Figure 5: Percentages of 'different' responses in the AB order (solid line: neutral base stimuli, dashed line: contrastive base stimuli; upper panel/Fig. 5a: sentence; lower panel/Fig. 5b: object).



As in previous studies [8], [14], we detect an order effect. Discrimination of stimuli in the BA order was inconsistent; the discrimination rate averaged 20% across the conditions. Despite these relatively low rates, the discrimination rate of the AA order was even lower, hardly exceeding 10%. This means that the listeners had no problem in recognizing these stimuli as identical.

4. DISCUSSION

Both identification rate and discrimination rate provide evidence for the fact that speakers make categorical use of F0 scaling differences in perceiving the different statements. As for identification, the scores show the expected S-shape. We interpret the weaker S-shape of the functions in the first experiment (cf. Fig. 3) as a hint for the fact that the listeners, when classifying the items as fitting in either the contrastive or the neutral context, also rely on prosodic cues other than F0 (i.e. intensity, sonority) that are more salient in the material preceding the (tonally marked) object *el diario*. The results from the identification test match with the both SE and RT data. As for SE, in line with [14], we interpret the low scores for the two 167Hz and 242Hz stimuli as a signal that the listeners agreed in their responses. The high scores for the 197Hz and 212Hz stimuli, on the other hand, signal a disagreement in the listeners' judgments. The SE value for the 197Hz stimulus is unexpectedly low (Fig. 3a), but the listeners' difficulties in qualifying this stimulus is clearly reflected in the high RT value (cf. Fig. 4a): Even though speakers agree in their responses, they need more time to make their decision. As for RT, we interpret the fact of the mean RT being shorter for the within-category contrastive-based stimuli than for the across-category stimuli as crucial evidence of a categorical difference, in line with [2]. We assume for the different patterns of the neutral-based stimuli that the long RTs to the two stimuli of 227Hz and 242Hz in Fig. 4b are effects of non-tonal cues. The tonal movement clearly signals an emphatic/contrastive nuance and the listeners thus classify the stimuli as being contrastive (cf. Fig. 3a). Nevertheless, these stimuli can hardly qualify as perfect candidates for expressing contrast, since they lack the non-tonal cues generally associated with emphasis/contrastiveness (intensity, sonority), in contrast to the test items stemming from the original contrastive stimuli. Listeners consequently need more time to classify them.

Evidence of the categorical contrast is also provided by the discrimination task. Listeners perceive (AB) differences between pairs of stimuli primarily in the crossover region of the identification curve – even though the peak does not exactly correspond with the category shift or RT peak, which is at 197Hz. The evidence is thus weaker, but nonetheless in line with our results from the identification task, SE, and RT. For example, [2] argues that the absence of a discrimination peak (as in [8]) does not automatically discount the claim of categorical distinction, provided that there is a clear RT peak (for difficulties in applying discrimination tasks to F0 cf. [2], [10], [11] and references cited therein).

In sum, the contrast between L* and L+H*+L is expressed only by scaling, i.e. by adding an H tone to a practically constant low environment. It is not expressed by alignment, i.e. there is no change in the phonological association with an L or H elbow.

Let's turn back to the heritage speakers (HS), who encountered serious problems in both the discrimination and the identification task (resulting in a zigzag curve) and thus seem to have deficiencies in their tonal system. In general, these two speakers classified all contrastive-based stimuli as rather contrastive, with all neutral-based stimuli being judged as rather neutral. It thus seems that they "heard" both the contrastive and the neutral nuances irrespective of the actual tonal shape. This might indicate that HS rely more on cross-linguistic non-tonal cues than on the tonal inventory, which is language-specific. The results from the HS thus speak in favor of the general applicability of the CP paradigm for intonational contrasts (cf. discussion in [5]), with a clear difference existing between the HS and the native speakers.

5. CONCLUSION

The present study shows that the scaling contrast between L+H*+L and L* in the tonal system of *Porteño*-Spanish is categorical. The results that were less clear can be understood in the context of well-known effects (e.g. the order effect) or can be explained by taking non-tonal cues into account.

6. ACKNOWLEDGMENTS

Research for this article was supported by the German Research Foundation (DFG). We would like to thank Oliver Niebuhr, Tomke Brüggemann, and Jasmina Živković for their essential help. All errors remain our own.

7. REFERENCES

[1] Boersma, P., Weenink, D. 2011. Praat: Doing phonetics by computer. *http://www.praat.org/*

[2] Chen, A. 2003. RT as an indicator of discrete intonational contrasts in English. *Proc. Eurospeech* 97-100.

[3] Forster, K.I., Forster, J.C. 2003. DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, Computers* 35, 116-124.

[4] Gabriel, C., Feldhausen, I., Peškova, A., Colantoni, L., Lee, S., Arana, V., Labastia, L. 2010. Argentinian Spanish intonation. In Prieto, P., Roseano P. (eds.), *Transcription of Intonation of the Spanish Language*. München: Lincom, 285-317.

[5] Gussenhoven, C. 2004. *The Phonology of Tone and Intonation*. Cambridge: CUP.

[6] Kohler, K. 1987. Categorical pitch perception. *Proc. 11th ICPhS* Tallinn, 331-333.

[7] Kohler, K., Niebuhr, O. 2007. The phonetics of emphasis. *Proc. 16th ICPhS* Saarbrücken, 2145-2148.

[8] Ladd, D.R., Morton, R. 1997. The perception of intonational emphasis: continuous or categorical? *J.Phon.* 25, 313-342.

[9] Liberman, A.M., Harris, K.S., Hoffman, H.S., Griffith, B.C. 1957. The discrimination of speech sounds within and across phoneme boundaries. *J. Exp. Psychology* 54, 358-368.

[10] Niebuhr, O., Kohler, K. 2004. Perception and cognitive processing of tonal alignment in German. *Proc. Int. Symp. Tonal Aspects of Languages.* Beijing, China, 155-158.

[11] Prieto, P. 2011. To be appeared. Experimental methods and paradigms for prosodic analysis. Ms. ICREA-UPF. In Cohn, et al. *Handbook of Laboratory Phonology*. OUP.

[12] Remijsen, B., van Heuven, V. 1999. Categorical pitch dimensions in Dutch: Diagnostic test. *Proc. 14th ICPhS* San Francisco, 1865-1868.

[13] Schneider, K., Lintfert, B., Dogil, G., Möbius, B. 2006. Phonetic grounding of prosodic categories. In Sudhoff, S., et al. (eds.), *Methods in Empirical Prosody Research*. Berlin: De Gruyter, 335-361.

[14] Vanrell, M. 2007. A tonal scaling contrast in Majorcan Catalan interrogatives. *J. of Port. Ling* 6, 147-178.