PROSODIC STRUCTURE AND PHONETIC PROCESSING: A CROSS-LINGUISTIC STUDY

Christophe Pallier[1,2], Anne Cutler[1] and Núria Sebastián-Gallés[2]
[1]Max-Planck-Institute for Psycholinguistics, Nijmegen, The Netherlands.
[2]Departament de Psicologia Bàsica, Universitat de Barcelona, Spain.
E-mail : pallier@lscp.ehess.fr

ABSTRACT

Dutch and Spanish differ in how predictable the stress pattern is as a function of the segmental content: it is correlated with syllable weight in Dutch but not in Spanish. In the present study, two experiments were run to compare the abilities of Dutch and Spanish speakers to separately process segmental and stress information. It was predicted that the Spanish speakers would have more difficulty focusing on the segments and ignoring the stress pattern than the Dutch speakers. The task was a speeded classification task on CVCV syllables, with blocks of trials in which the stress pattern could vary versus blocks in which it was fixed. First, we found interference due to stress variability in both languages, suggesting that the processing of segmental information cannot be performed independently of stress. Second, the effect was larger for Spanish than for Dutch, suggesting that that the degree of interference from stress variation may be partially mitigated by the predictability of stress placement in the language.

1. INTRODUCTION

The prosodic and segmental dimensions of speech are fully integrated in the speech signal. Any prosodic contrast between two signals is realised via changes in how the segments are uttered — with longer or shorter duration, greater or lesser amplitude, more or less pitch movement, for example. Yet recent studies from a number of different languages have suggested that the processing of prosodic structure in the form of stress is to some extent independent of the processing of segmental structure. Thus lexical access by English listeners does not appear to take into account stress distinctions, of the kind that differentiate between word pairs such as *'trusty* and *trus'tee* [1]; similar results with such word pairs also appear in Dutch [2]. French listeners ignore accent differences when making similarity judgments about nonwords [3].

The selective attention paradigm [4] offers a powerful measure of whether two dimensions of an auditory signal may or may not be processed independently. In this task listeners categorise stimuli on the basis of one auditory dimension; the independent variable is the presence versus absence of variation in another dimension, notionally irrelevant to the categorisation task. With simple CV syllables, classification of segments is slowed by irrelevant variation in the syllable's pitch [5-8], suggesting that the processing of segmental information and of at least one type of prosodic information are not independent. In the present study we used the selective attention task to investigate the processing of stress and segmental structure in CVCV bisyllables. If the results from simple pitch variation are representative of all types of prosodic information, then stress variation will also interfere with segmental decisions; but if higher levels of prosodic structure such as stress (which requires a polysyllabic domain for contrasts to be displayed) are indeed processed independently of segmental structure, as the results summarised above seem to suggest, then we will observe no such interference.

We carried out the comparison in two languages, which both exhibit stress contrasts, but differ in the degree of dependence between the prosodic and segmental structure of words. In Dutch, stress contrasts are usually accompanied by contrasts in syllable weight, whereby stress is assigned to the heavier syllable. In Spanish, bisyllables in which the two syllables have equal weight are much more common. Since the placement of Dutch stress is thus generally more predictable than the placement of Spanish stress, we may expect that segmental judgments are more likely to be independent of stress variation in Dutch than in Spanish.

2. EXPERIMENT 1: SPANISH SUBJECTS

2.1 Method

2.1.1 Materials

The pseudowords *deki* and *nusa* were chosen as stimuli for the experiment; both would make phonologically acceptable lexical items in Spanish or in Dutch. One Spanish speaker and one Dutch speaker recorded two tokens of both *deki* and *nusa*: one with the stress falling on the first syllable, and the other with the stress on the second syllable. Thus, there were four Spanish items and four Dutch items (*'deki, 'nusa, de'ki* and *nu'sa*). These stimuli were digitized at 16 Khz and stored on the hard drive of a PC computer.

2.1.2 Procedure

Each subject was tested in an individual booth, facing a PC that controlled the presentation of the stimuli and measured the response times, using the software EXPE [9].

The task was speeded classification. Each trial started with the auditory presentation of a stimulus which the subject had to identify. She indicated her response by pressing one of two buttons corresponding respectively to *deki* and *nusa* (irrespective of the stress position). She had a 1500 msec time window (starting at stimulus onset) to do this, after which feedback was displayed for 800 msec: if the decision was correct, the reaction time from stimulus' onset was given, whereas in case of an incorrect response (or no response), the mapping of stimuli to response keys was presented on the screen. The screen was then cleared for one second, and the next trial started.

The experiment consisted of blocks of 64 trials. There were two types of blocks: "Varied" blocks where the four items (*'deki, 'nusa, de'ki, nu'sa*) appeared and where the stress pattern could thus vary from trial to trial; and "Fixed" blocks where only two items (e.g. *'deki* or *'nusa*) were used as stimuli and where the stress pattern was constant. All the items in a block belonged to the same language, Spanish or Dutch.

After a practice block with sixteen trials and four stimuli, each subject was presented with four experimental blocks: two "varied" and two "fixed" blocks. Among the "fixed" blocks, one contained the items with initial stress while the other contained the items with final stress. The order of blocks was counterbalanced across subjects and a given subject heard only stimuli from one language (Spanish or Dutch). In other words, half the subjects were assigned to the Spanish material and the other half to the Dutch material. The order of items in a block was randomized for each subject.

2.1.3 Subjects

Sixty-four native speakers of Spanish, all students from the university of Barcelona, participated in the experiment in exchange for course credits.

2.2 Results

Mean response times and error rates are displayed in Table 1. An analysis of variance on the response times, which crossed the factors Stress (Varied vs. Fixed) and Language of the stimuli (Spanish vs. Dutch), revealed a significant effect of Stress (F(1,56)=64; p<.0001) which interacted with the Language of the stimuli (F(1,56)=8.8; p<.01). One-way ANOVAs of the Spanish and Dutch stimuli separately showed that the effect of Stress was nevertheless significant in both cases (Spanish: F(1,28)=14; p<.001; Dutch: F(1,28)=54; p<.0001). A similar ANOVA on error rates yielded no significant effect.

Table 1: Mean classification times (in ms) and error rates of Spanish subjects.

	Stress pattern		
	Fixed	Varied	*Difference*
Stimuli			
Spanish	432 / 3.0%	447 / 2.7%	15 / -0.3%
Dutch	444 / 2.7%	476 / 3.1%	32 / +0.4%

Spanish subjects clearly show the interference effect; that is, they are significantly slowed down by the stress variability. The effect is larger for the stimuli pronounced by a foreign speaker.

3. EXPERIMENT 2: DUTCH SUBJECTS

Stress is more predictable in Dutch than in Spanish. Will Dutch subjects be better able than the Spanish to ignore the stress variations and focus on the segments? To investigate this question, we replicate the previous experiment with Dutch subjects.

3.1 Method

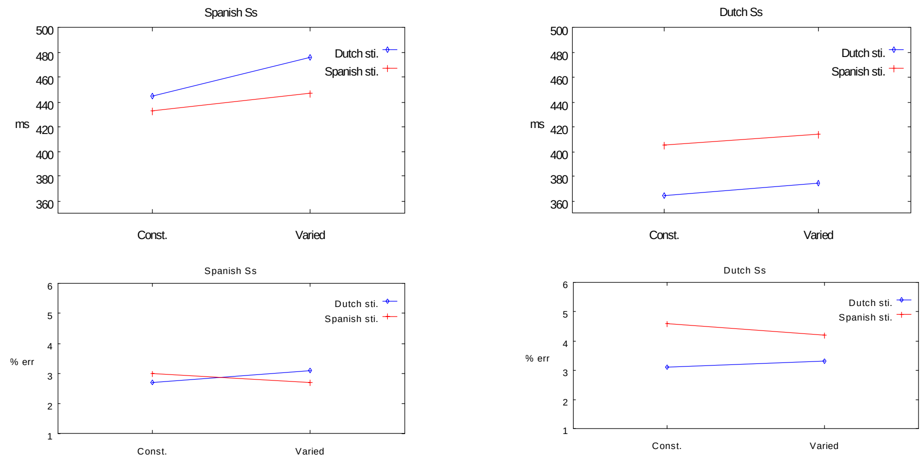
3.1.1 Material & Procedure

The same materials, procedure and design as in experiment 1 were employed.

3.1.2 Subjects

Fifty-six subjects from the pool of subjects of the Max-Planck-Institute for Psycholinguistics, in Nijmegen, participated in the experiments.

Figure 1: Mean classification times and error rates of Spanish and Dutch subjects



3.2 Results

Mean response times and error rates are displayed in Table 2. An analysis of variance on the response times, which crossed the factors Stress (Varied vs. Fixed) and Language of the stimuli (Spanish vs. Dutch), revealed a significant effect of Stress ($F(1,48)=11$; $p<.01$) and of Language of the stimuli ($F(1,48)=12$; $p<.01$), with no interaction between them. T-tests restricted to Spanish and Dutch stimuli respectively showed that the effect of Stress was significant for the Dutch stimuli ($F(1,24)=10$; $p<.01$) and marginally significant for Spanish stimuli ($F(1,24)=3.55$; $p=.07$). A similar ANOVA on error rates yielded a significant effect due to Language of Stimuli ($F(1,48)=4.53$; $p<.05$).

Table 2: Mean classification times (in ms) and error rates of Dutch subjects.

	Stress pattern		
	Fixed	Varied	*Difference*
Stimuli			
Spanish	405 / 4.6%	414 / 4.2%	9 / -0.4%
Dutch	365 / 3.1%	375 / 3.3%	10 / +0.2%

Dutch subjects display interference effects of the same size whatever the language of the stimuli, though they are faster and make less errors when the stimuli are pronounced in their maternal language.

3.3 Combined Analysis

The results from the two experiments are displayed in Figure 1. In order to compare them, we combined both sets of data and performed an ANOVA in which the factor "Language of the subject" was added. The RTs in the fixed condition (413 ms) were significantly faster than in the varied condition (430 ms; $F(1, 104) = 67.24$, $p < .001$). As predicted, this interference effect was larger for the Spanish subjects (23 ms) than for the Dutch subjects (9 ms); however, it was only significantly larger for the Spanish subjects with the Dutch stimuli, and interference was observed for both subject groups and both materials sets. There was no statistically significant difference between response time to tokens with first- versus second-syllable stress. RTs of the Dutch subjects were significantly faster (390 ms) than those of the Spanish subjects (450 ms); however, this reflected a speed-accuracy tradeoff, since the error analysis showed a significant difference between subject groups: Spanish subjects made significantly fewer errors (2.9%) than Dutch subjects (3.8%). No other effect reached significance in the error analysis. Finally, despite the fact that each subject only heard (multiple repetitions of) four tokens, responses were facilitated when the tokens had been spoken by a native speaker of the subject's language (405 ms in comparison with 435 ms for non-native tokens).

4. CONCLUSION

These experiments have produced a clear result: listeners cannot ignore stress information while making a segmentally-based classification. This is true both in Spanish, in which stress cannot easily be predicted from syllable weight, and in Dutch, in which stress is more predictable.

The findings have potential implications for the relevance of prosodic structure to phonetic processing in speech recognition. As described in the introduction, previous research suggests that French listeners ignore accent information [3], and English listeners do not use stress distinctions in lexical access [1]. The French result presumably reflects the irrelevance of accent distinctions to the determination of word identity in that language. The English result, however, may also reflect language-specific distributional effects. Recent evidence from perceptual studies in a number of languages suggests that stress (and indeed other types of prosodic structure) may be exploited in spoken-word recognition where it significantly contributes to reducing the number of potential word candidates activated by a given input (see [10] for a review). In English, stress is nearly always unambiguously signaled by segmental structure (vowel quality), so that suprasegmental correlates of stress contribute little to reduction of the pool of word candidates; but in Dutch and in Spanish, the suprasegmental structure can be more informative. Indeed, Koster and Cutler [11] report that mis-stressing Dutch words by varying the suprasegmental pattern alone has a strong adverse effect on the efficiency with which a word can be recognised, a result which contrasts with earlier reports from English (e.g. [12]).

The generality of the interference effect in the present study suggests that the processing of segmental information cannot be performed independently of stress. Nevertheless, the fact that the effect was larger for speakers of Spanish than for speakers of Dutch suggests that the degree of interference from stress variation may be partially mitigated by the predictability of stress placement in the language.

5. ACKNOWLEDGMENTS

This work was supported by a grant from the Spanish Ministry of Education ("Acciónes Integradas") and part of it was realised when Christophe Pallier was supported by a post-doctoral grant from the Fyssen Foundation. We thank Mariëtte Koster and Mark Scholten for technical assistance.

6. REFERENCES

[1] A. Cutler, "*Forbear* is a homophone: lexical prosody does not constrain lexical access," *Language and Speech*, vol. 29, pp. 201-220, 1986.

[2] W. Jongenburger and V. J. van Heuven, "The role of linguistic stress in the time course of word recognition in stress-accent languages," presented at 4th European Conference on Speech Communication and Technology, 1995.

[3] E. Dupoux, C. Pallier, N. Sebastian, and J. Mehler, "A destressing "deafness" in French?" *Journal of Memory and Language*, vol. 36, pp. 406-421, 1997.

[4] W. R. Garner, *The Processing of Information and Structure*. Potomac Md: Erlbaum, 1974.

[5] C. C. Wood, "Parallel processing of auditory and phonetic information in speech discrimination," *Perception & Psychophysics*, vol. 15, pp. 501-508, 1974.

[6] L. Lee and H. C. Nusbaum, "Processing interactions between segmental and suprasegmental information in native speakers of English and Mandarin Chinese," *Perception & Psychophysic*, vol. 53, pp. 157-165, 1993.

[7] B. Repp and H.-B. Lin, "Integration of tonal information in speech perception: A cross-linguistic study," *Journal of Phonetics*, vol. 18, pp. 481-495, 1990.

[8] J. L. Miller, "Interactions in processing segmental and suprasegmental features of speech," *Perception & Psychophysics*, vol. 24, pp. 175-180, 1978.

[9] C. Pallier and E. Dupoux, "EXPE: an expandable programming language for on-line psychological experiments," *Behavior Research Methods, Instruments, & Computers (in press)*, 1997.

[10] A. Cutler, D. Dahan, and W. v. Donselaar, "Prosody in the comprehension of spoken language: A literature review," *Language and Speech*, vol. 40.

[11] M. Koster and A. Cutler, "Segmental and suprasegmental contributions to spoken-word recognition in Dutch," presented at Eurospeech '97: 5th European Conference on Speech Communication and Technology, Rhodes, Greece, 1997.

[12] L. M. Slowiaczek, "Effects of lexical stress in auditory word recognition," *Language and Speech*, vol. 33, pp. 47-68, 1990.