The use of the Odd-One-Out task in the study of the perception of lexical stress in Spanish by German-speaking listeners?

Schwab, Sandra ; Dellwo, Volker

The use of the Odd-One-Out task in the study of the perception of lexical stress in Spanish by German-speaking listeners

Sandra Schwab[1, 2], *Volker Dellwo*[1]

[1] Institut für Vergleichende Sprachwissenschaft, Universität Zürich, Switzerland
[2] Ecole de langue et de civilisation françaises, Université de Genève, Switzerland
sandra.schwab@uzh.ch, volker.dellwo@uzh.ch

Abstract

In the present research we investigated the perception of Spanish stress in German-speaking listeners in comparison with native Spanish listeners. We used a cognitively demanding Odd-One-Out task and stimuli with variability in voice and/or in intonation. The main findings showed that the German-speaking listeners were able to perceive the Spanish lexical stress to a very high degree (76% of correct responses), but that their performance was lower than the Spanish listeners' performance (90%). The difference between German and Spanish speakers was mainly due to the German speakers' poorer detection of the *odd* in two specific accentual contrasts. The implications on the *stress deafness* hypothesis are discussed.

Index Terms: stress deafness, German, Spanish, Odd-One-Out task.

1. Introduction

While speakers of some languages can easily identify the position of lexical stress in L2, speakers of other languages show some difficulties in performing this task. The difficulty in perceiving or discriminating accentual contrasts that do not exist in L1 is the basis of the *stress deafness* hypothesis (e.g., [1]-[3].). According to this hypothesis, the degree of *stress deafness* is related to the accentual properties of L1, more specifically to the nature of lexical stress (free or fixed). In a free-stress language such as Spanish, German or English, lexical stress has a distinctive function, since it distinguishes segmentally identical words such as in Spanish *<u>nú</u>mero* (['numeɾo], engl. *(the) number*) and *nu<u>me</u>ro* ([nu'meɾo], engl. *I number*) or in German *<u>um</u>fahren* (['ʊmfarən], engl. *run into*) and *um<u>fa</u>hren* ([ʊm'farən], engl. *drive round*)[1]. As a consequence, speakers of a free-stress language encode the accentual information in their mental representation of the words. On the other hand, the position of stress in a fixed-stress language such as French or Finnish is not variable, and thus not contrastive. Consequently, the accentual information does not need to be stored in the lexical representation of speakers of a fixed-stress language. The *stress deafness* hypothesis claims that speakers of fixed-stress languages have more difficulties in perceiving stress contrasts in L2 than speakers of free-stress languages, since they are not able to encode the accentual information in their mental lexicon (e.g., [3]). In view of this, speakers of German should not show evidence of *stress deafness* in a L2 like Spanish. Alternatively, since the acoustic correlates of stress are to some degree language-specific, it is unclear whether German listeners show the same performance in stress distinction compared to Spanish listeners.

In addition, the degree of *stress deafness* has been shown to depend on other factors like the cognitive load involved in the task and the variability in voice (i.e., speaker) ([1]-[3]). Among the test paradigms with a low cognitive load, we find, for example, the identification task, in which the participants are asked to identify what they perceived in an open or closed set of responses (used for example in [5]-[6]). Another paradigm with a low cognitive load is the AX discrimination task, where the participants have to decide whether two stimuli are the same or different (used in [1] for example). Among the paradigms with a higher cognitive load, we find, for example, the ABX discrimination task, in which the participants hear three stimuli and are instructed to indicate whether the third stimulus (X) is the same as the first stimulus (A) or as the second stimulus (B) (used for example in [1]). The ABX task presents two major problems. First, since the participants, in order to make their decision, have to hold Stimulus A in the short-term memory while they hear stimulus B, there is a bias towards Stimulus B, as it is more current in the short-term memory (e.g., [7]). Second, it seems conceivable that the participants perform the ABX task by simply judging whether Stimulus B is the same or different from Stimulus X. In that case, the ABX task converts into an AX discrimination task. These disadvantages led [7] to use a variation of the ABX task, namely an AXB task, where Stimulus X is presented between Stimulus A and Stimulus B. [2] developed another cognitively demanding paradigm, namely the sequence-recall task. In this task, two pseudowords differing in stress position (e.g., *<u>pí</u>ki* with stress on the first syllable, and *pi<u>kí</u>* with stress on the second syllable) are associated with two keys on a keyboard (e.g., *<u>pí</u>ki* = key 1 and *pi<u>kí</u>* = key 2). The participants hear sequences composed of the two pseudowords (e.g., *<u>pí</u>ki*, *<u>pí</u>ki*, *pi<u>kí</u>*, *<u>pí</u>ki*) and have to reproduce them with the corresponding keys (e.g., 1, 1, 2, 1).

In the present research, we added a new level of complexity in the task for the examination of *stress deafness* by using the Odd-One-Out task ([8]). In this task, the participants hear a sequence of three segmentally identical stimuli (e.g., *numero*). Among them, two stimuli present the same accentual pattern (e.g., stress on the penultimate syllable) and one (i.e., the *odd*) presents a different accentual pattern (e.g., stress on the final syllable). The participants' task is to indicate which stimulus is the *odd* element of the sequence. As can be seen, this task lies between the AXB task and the

[1] The underlined syllable in these examples, and in the rest of the paper, corresponds to the stressed syllable.

sequence-recall task. We argue, however, that in comparison with the AXB task, the Odd-One-Out task constitutes a more real-life situation, since listeners are typically not primed about the stress pattern of a word in advance (unless the situational context allows this). The Odd-One-out task is thus more cognitively demanding, since the target stimulus (i.e., the *odd*) is not always in the middle position within the sequence, but also in the first and third positions. Unlike the sequence-recall task, Odd-One-Out presents the advantage to provide the possibility to collect not only the participants' correct/incorrect responses, but also their reaction times, which are relevant in investigating *stress deafness* ([1], [3], [6], [7]).

In summary, in the present research we tested the ability of German-speaking listeners without knowledge of Spanish to identify stress in Spanish words based on the Odd-One-Out paradigm.

2. Method

2.1. Participants

Two groups of listeners participated in this experiment: 1) 17 native speakers of Spanish from Barcelona (hereafter "Spanish listeners"; mean age: 23.5 years, stdev: 6); 2) 22 native speakers of German from Zurich (hereafter "German-speaking listeners"; mean age: 24.5 years, stdev: 3.4). The German-speaking participants had no knowledge of Spanish, Italian or Portuguese (i.e., romance languages with free-stress). Participants were students of Universitat Autònoma de Barcelona (Spanish listeners), Universitat Pompeu Fabra (Spanish listeners) and Universität Zürich (German-speaking listeners). Listeners were paid for their participation.

2.2. Material

Two triplets of trisyllabic Spanish words were used. The three elements of each triplet differed with respect to the stressed syllable. Each triplet was composed of a proparoxytone word (i.e., first syllable stressed word) (hereafter "PP"; *número*, *válido*), a paroxytone word (i.e., penultimate syllable stressed word) (hereafter "P"; *numero*, *valido*) and an oxytone word (i.e., final syllable stressed word) (hereafter "O"; *numeró*, *validó*). Two native female speakers of Peninsular Spanish (Speaker 1 and Speaker 2) produced twice the six words in a declarative sentence with a falling intonation (i.e., *Le dijo a Pat "número"*; engl. *He told Pat "número".*) and in an interrogative sentence with a rising intonation (i.e., *¿Le dijo a Pat "número"?*; engl. *Does he tell Pat "número"*?). For each sentence, we extracted the last word (e.g., *número*) and we created sequences composed of three segmentally identical words, separated by 500 ms (e.g., *número-numero-número*). The sequences were composed of two words with the same accentual pattern (e.g., PP; *número*) and of one word with a different accentual pattern (e.g., P, *numero*). The word with the different accentual pattern was the *odd* element within the sequence and constituted thus the target word. Among all the words (6 words x 2 speakers x 2 intonations), only the 6 words produced by Speaker 1 with the falling intonation were used as target words, while the other words were only used to introduce variability within the sequences.

We constructed the sequences in such a way that a large phonetic variability was present (as recommended in [1]-[3]). Stimuli were created in the following way: 1) The accentual pattern of the target word was counterbalanced (i.e., same number of PP, P and O words as being target words); 2) All the accentual contrasts were tested (i.e., PP target word paired with P words or with O words; P target word paired with PP words or with O words; O target word paired with PP words or with P words); 3) The position of the target word within the sequence was counterbalanced (i.e., same number of target words in position 1, 2 and 3); 4) Half of the sequences was composed of words produced by one speaker, and half of them was composed of words produced by two speakers. This was done to introduce phonetic variability induced by a variation in voice (i.e., speakers); 5) Half of the sequences was composed of words produced with the same intonation pattern, and half of them was composed of words produced with two intonation patterns. Phonetic variability was added in that case with a variation in intonation.

A total of 144 test sequences was used in the experiment. We also introduced 72 filler sequences in which the target word was produced by Speaker 2 and/or with a rising intonation to avoid the participants to develop a strategy that enabled them to identify the *odd* element of the sequence. These filler sequences were not included in the analyses.

2.3. Procedure

The participants performed an Odd-One-Out task. After hearing each sequence, they had to indicate, as fast as possible which of the three elements was the *odd* one (i.e., the different one), by pressing the corresponding key (1, 2 or 3) on a response box. They were told that the *odd* element differed with respect to the stressed syllable. They had 3 seconds to answer and the next sequence was presented (even if the participants did not give any answer). Each participant received a different randomization of the 216 sequences. The experiment lasted 25 minutes.

2.4. Data analysis

One German-speaking participant had to be excluded from the analysis due to a particularly large number of missing data (43%), which led to 21 German-speaking participants (i.e., non-natives) and to 17 Spanish participants (i.e., natives)[1].

Statistical analyses were carried out with the R software (version 3.1.3; R Development Core Team, 2014, *lmerTest* package, [4]). We ran two analyses on the correct/incorrect responses using mixed-effects logistic regression models [9] (RTs will be examined in future research).

The first analysis, which aimed at validating the use of the Odd-One-Out task in the examination of the perception of lexical stress, was performed on the native Spanish listeners' data only. In this analysis, the following predictors were entered into the model: Word (*valido, numero*) and *Odd* position (1, 2, 3), interaction Word x *Odd* Position, and Presentation order[2].

[1] Although the global rate of missing data was relatively low (2.54%), the German-speaking listeners presented more missing data (3.14%, stdev: 4.22) than the Spanish listeners (1.8%, stdev: 2.7; $\chi^2(1) = 9.9$, $p < .01$).

[2] The nominal variable "Word" (*valido, numero*) was recoded into a [1, -1] dummy variable, the nominal variable "*Odd* position" (1, 2, 3) was recoded into [0, 1] dummy variables, and the numerical variable "Presentation order" was rescaled for the values being between 0 and 1.

In the second analysis, we examined the effect of L1 and accentual contrast on the participants' accuracy in detecting the *odd*. In this model, the predictors were L1 (Spanish, German), Accentual contrast (PP paired with P, PP paired with O, etc.) and the interaction L1 x Accentual Contrast. The following controlled variables were also entered into the model: Word (*valido, numero*) and *Odd* position (1, 2, 3), Presentation order and the interaction between each of the controlled variables and L1[1]. In this analysis, the non-significant controlled variables or interactions were removed from the final model.

In both analyses, participants and items were entered as random variables. The significance of the main effects and interactions was assessed with likelihood ratio tests that compared the model with the main effect or interaction to a model without it. For clarity's sake, the results and figures are presented in percentages, although all statistical analyses have been performed on raw data (correct/incorrect responses).

3. Results and discussion

3.1. Validation of the Odd-One-Out task in native Spanish listeners

Spanish listeners present a performance of 90% of correct responses, which suggests that the identification of the *odd* does not constitute a difficult task for native listeners (chance level was 33.33%). Moreover, we observed no effects of Word (*numero* = 91.18% and *valido* = 90.43%; $\chi^2(1) = 0.88$, n.s.), *Odd* position (1 = 92.38%, 2 = 89.34%, 3 = 90.69%; $\chi^2(2) = 3.16$, n.s.) or interaction between Word and *Odd* position ($\chi^2(2) = 1.23$, n.s.). However, an effect of Presentation order is present ($\chi^2(1) = 7.72$, $p < .01$). The error rate slightly increased along the experiment ($\beta = 0.79$, $SE = 0.27$, $z = 2.93$), which seems not to be so surprising in a 30-minute experiment.

Importantly, these findings let us conclude that the Spanish listeners' responses were not biased by the word (*numero* or *valido*) or by the position of the odd in the sequence, which validates the use of the Odd-One-Out task in the perception of lexical stress.

[1] The nominal variable "L1" (Spanish, German) was recoded into a [-1, 1] dummy variable, the nominal variable "Accentual contrast" (PP paired with P, PP paired with O, etc.) was recoded into [0, 1] dummy variables. Moreover, the nominal variable "Word" (*valido, numero*) was recoded into a [1, -1] dummy variable, the nominal variable "*Odd* position" (1, 2, 3) was recoded into [0, 1] dummy variables, and the numerical variable "Presentation order" was rescaled for the values being between 0 and 1.

3.2. Effect of L1 and accentual contrast in the detection of the *odd*[2]

Figure 1 presents the percent correct identification of the *odd* as a function of L1 (Spanish and German). As can be seen, the performance of the native Spanish listeners is higher (90% of correct responses) than the performance of the German-speaking listeners (76% of correct responses) ($\chi^2(1) = 29.39$, $p < .001$). Despite the presence of some outliers (indicated by a star in Figure 1), especially in the German-speaking listeners, their performance is above chance level (Binomial values between 0.000 and 0.005, for $p = 0.33$). These findings indicate that the speakers of German were able to detect the accentual *odd* element within a sequence in Spanish, but they were not as good as the native speakers of Spanish[3].

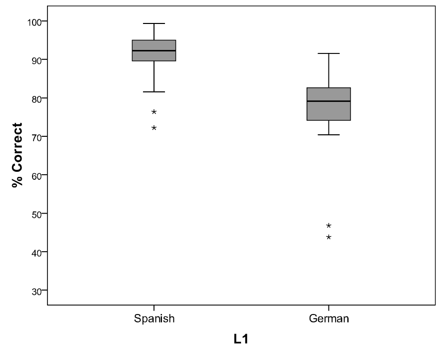

Figure 1: *Percent correct identification of the "odd" as a function of L1 (Spanish and German).*

Figure 2 presents the percent correct identification of the *odd* as a function of L1 and Accentual contrast. In this figure, "PP paired with P", for example, means that the *odd*, which is proparoxytone (i.e., with stress on the first syllable), appeared in the sequence among paroxytone words (i.e., with stress on the penultimate syllable). Along the same line, in "O paired with P", the *odd* is oxytone (i.e., with stress on the final syllable) and was among paroxytone words (i.e., with stress on the penultimate syllable).

As can be seen, we observe an effect of Accentual contrast ($\chi^2(5) = 20.74$, $p < .001$), and more interestingly an interaction between L1 and Accentual contrast ($\chi^2(5) = 133.53$, $p < .001$). Spanish and German-speaking listeners are not equally

[2] As far as the controlled variables are concerned, the effect of *Odd* position was significant ($\chi^2(2) = 11.28$, $p < .001$): the *odd* was harder to detect in Position 2. Moreover, the *odd* was better identified in the *numero* sequences than in the *valido* sequences ($\chi^2(1) = 7.43$, $p < .01$), especially in the German-speaking participants ($\chi^2(1) = 16.49$, $p < .001$). The effect of Presentation order was different in Spanish- and German-speaking participants ($\chi^2(1) = 22.64$, $p < .001$): while the error rate increased along the experiment in Spanish-speaking participants (as previously mentioned), it decreases in German-speaking participants.

[3] We ran the same analysis without the two German-speaking outliers to ensure that the L1 effect was not only due to these two participants. The results without the two outliers were similar to the results with the outliers (effect of L1: $\chi^2(1) = 29.76$, $p < .001$; effect of accentual contrast: $\chi^2(5) = 21.86$, $p < .001$; interaction L1 x Accentual contrast: $\chi^2(5) = 138.20$, $p < .001$).

sensitive to the accentual contrasts. Post-hoc analyses (with Tukey corrections) showed that the Spanish listeners have more difficulties with paroxytone *odds* (i.e., with stress on penultimate syllable), namely with "P paired with PP" and "P paired with O" (especially in comparison with "O paired with P"). The German-speaking listeners present a lower performance in "PP paired with O" and "O paired with PP". Thus, German-speaking listeners seem to have more difficulties with the pairing of proparoxytone words (i.e., stress on the first syllable) and oxytone words (i.e., stress on the final syllable). Interestingly, it is in these cases that we observe significant differences between the Spanish and the German-speaking listeners (in addition to "O paired with P").

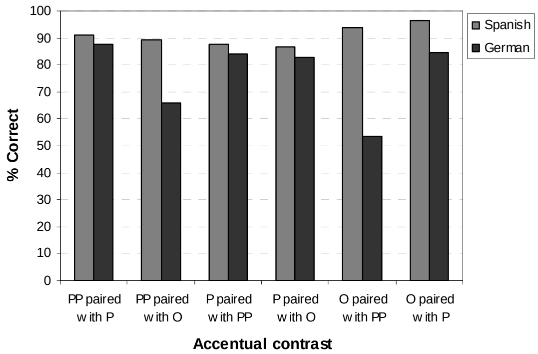

Figure 2: *Percent correct identification of the "odd" as a function of Accentual contrast and L1 (Spanish and German).*

4. General discussion

The results of this research are interesting in several respects. First, they revealed that the Odd-One-Out task constitutes an appropriate task in the examination of the perception of lexical stress. Second, they showed that the German-speaking listeners are able to perceive the Spanish lexical stress to a very high degree (76% of correct responses), but that their performance is lower than the Spanish listeners' performance.

This finding supports the view that the correct perception of stress might be possible in an unfamiliar language, but that the performance is poorer. It is possible that this is related to the realization of stress in Spanish which might be not exactly the same as in German. [10], for example, showed that speakers of a language with vowel reduction such as English (or German) might present difficulties in hearing the position of stress in languages with less vowel reduction (like Spanish). Thus, it seems conceivable that, since languages vary in different ways in their stress correlates, listeners of some languages might be better at detecting stress in some languages than in others. This interpretation is particularly relevant for key experiments in the examination of the *stress deafness*, such as [1]. In these experiments, Spanish and French listeners judged stress in nonsense words produced by Dutch speakers. It seems possible that Dutch stress might in some respects be closer to Spanish than to French, which might contribute (besides the fact that stress is not distinctive in French) to a weaker performance of French listeners, and hence the belief that French listeners are deaf to stress.

Moreover, we have observed that the difference between German-speaking and Spanish listeners was mainly due the German-speaking listeners' poorer detection of the *odd* in accentual contrasts that paired oxytone and proparoxytone words (i.e., words with stress on the final and first syllable, respectively). This finding is surprising since the three accentual patterns (i.e., proparoxytone, paroxytone and oxytone) exist in the German language (e.g., [11]). Nevertheless, the German accentual pattern seems to be related to the syllabic structure of the word. For example, [12] examined a corpus of trisyllabic German words ending with different syllabic structures and showed that 58.3% of the words ending with a vowel (i.e., the syllabic structure of the words used in the present experiment) were paroxytone, 37.7% were proparoxytone, and only 4% were oxytone. Thus, the German-speaking listeners might have been disturbed by the oxytone stress of a word ending with a vowel. The poor detection of the *odd* in sequences of oxytone and proparoxytone words can also be explained by the fact that oxytone German words rarely contrast with proparoxytone or paroxytone words, whereas paroxytone and proparoxytone do (e.g., *umfahren* versus *umfahren*). On the other hand, surprisingly, the German-speaking listeners did not present particular difficulties in sequences where oxytone words are paired with paroxytone words. In that case, their perception of stress might have been influenced (and facilitated) by their expectation that the paroxytone pattern may constitute the default pattern in Spanish as it is the case in German ([11]).

Finally, we cannot exclude that the difference between Spanish and German-speaking listeners might be due to the fact that the former had access to the lexical representation of the words they heard, whereas the latter did not. To test this hypothesis, we ran the above experiment with five German-speaking listeners with knowledge of Spanish (between 3 and 6 years of Spanish courses). Results showed that they reached 84% of correct identification of the *odd*. Nevertheless, in spite of the higher performance in comparison with the German-speaking listeners without knowledge of Spanish, the participants with knowledge of Spanish still presented a lower performance in sequences where oxytone and proparoxytone words were paired (80.70% and 76.52 compared to 83.70%-88.70% for the other accentual contrasts). The findings of this pilot study suggest that the knowledge of the language does not enable the German-speaking listeners to overcome the difficulties related to the accentual pattern easily. Thus, it is not in support of the hypothesis according to which the absence of lexical access in the German-speaking listeners without knowledge of Spanish was responsible for their lower performance, compared to the Spanish listeners.

In conclusion, the present research validates the introduction of the Odd-One-Out paradigm in the examination of the perception of lexical stress. Moreover, it confirms that, although the perception of stress in L2 is somehow conditioned by the accentual characteristics of the L1, the speakers of a free-stress language like German do not experience particular difficulties in perceiving stress in another free-stress language like Spanish.

5. Acknowledgements

We would like to thank M. Amato, L. Baqué, Mª A. Barquero, M. Carranza, J. M. Lahoz, J. Llisterri, M. Machuca and A. Ríos for their helpul advice during the elaboration and analysis of this experiment. S. Schwab's work was supported by the Swiss National Foundation (grant Ambizione PZ00P1_148036/1).

6. References

[1] Dupoux, E., Pallier, C., Sebastián, N. and Mehler, J., "A destressing 'deafness' in French?", *Journal of Memory and Language*, 36(3), 406-421, 1997.

[2] Dupoux, E., Peperkamp, S. and Sebastián, N., "A robust method to study stress 'deafness'", *Journal of the Acoustical Society of America*, 110(3-1), 1606-1618, 2001.

[3] Dupoux, E., Sebastián, N., Navarrete, E. and Peperkamp, S., "Persistent stress 'deafness': The case of French learners of Spanish", *Cognition*, 106(2), 682-706, 2008.

[4] Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2014). lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-20. http://CRAN.R-project.org/package=lmerTest.

[5] Muñoz García, M. *La perception et la production de l'accent lexical de l'espagnol par des francophones: aspects phonétiques et psycholinguistiques*, Thèse de doctorat, Université de Toulouse, 2010.

[6] Schwab, S. and Llisterri, J., "The perception of Spanish lexical stress by French speakers: Stress identification and time cost", in Wrembel, M., Kul, M., Dziubalska-Kołacyk, K., [Eds.], *Achievements and perspectives in SLA of speech: New sounds 2010*. Vol. 1. Frankfurt am Main: Peter Lang, 229-242, 2011.

[7] Tremblay, A. "Phonetic variability and the variable perception of L2 word stress by French Canadian Listeners". International *Journal of Bilingualism*, 13, 35-62, 2009.

[8] Frearson, W. M. and Eysenck H. J. "Intelligence, Reaction Time (RT) and a New "Odd-Man-Out" RT Paradigm". *Personality and Individual Differences*, **7**(6), 807–817. 1986.

[9] Baayen, R.H., Davidson, D.J. and Bates, D.M., "Mixed effects modeling with crossed random effects for subjects and items", *Journal of Memory and Language*, 59, 390-412, 2008.

[10] Taylor, B. and Hellmuth S. "The effect of L1 background on perception of varying phonetic correlates of word accent" in *19th International Congress of Linguists. Geneva. July 2013*, 2013.

[11] Wiese, Richard. *The Phonology of Ger*man. Oxford: Clarendon Press, 1996/2000.

[12] Janssen, U. *Untersuchungen zum Wortakzent im Deutschen und Niederländischen*. PhD Dissertation. Universität Düsseldorf, 2003.