The role of stress perception in the assignment of written accent in Spanish

Sandra Schwab[1]*, Carla V. Jara Murillo*[2]

[1] ELCF, Université de Genève, Switzerland
[2] Escuela de Filología, Lingüística y Literatura, Universidad de Costa Rica, Costa Rica
Sandra.Schwab@unige.ch, Carla.Jara@ucr.ac.cr

Abstract

The aim of this investigation is to examine whether the adults' difficulty in placing the written accent in Spanish words is related to their ability in perceiving stress. The following variables were also taken into account in this study: the participant's education level (academic and non-academic), the stimulus lexical status (words and non-words), accentual pattern (proparoxytone, paroxytone and oxytone words) and length (2, 3 and 4 syllables). Participants performed a stress identification task and a word spelling task. Besides the effects of lexical status, education level and accentual pattern, results show an effect of the stress perception in the assignment of the written accent: stimuli with a correctly identified stress were more likely to be correctly written (i.e. with or without written accent) than the incorrectly perceived stimuli. This finding reinforces the idea that there is a relationship between prosodic and written skills.

Index Terms: stress perception, written accent, Spanish

1. Introduction

In Spanish, lexical stress is distinctive and therefore its placement is variable. As it is well known, stress distinguishes Spanish words such as *límite* (['limite] *limit*), *limite* ([li'mite] *(that) I limit*) and *limité* ([limi'te] *I limited*). Thus, it can appear on the last syllable of the word (oxytone word), on the penultimate syllable (paroxytone word) or on the antepenultimate syllable (proparoxytone word)[1]. However, the paroxytone pattern is by far the most general. It has been shown in [1] that in a corpus of 9219 Spanish words including propaxytone, paroxytone and oxytone words (excluding monosyllabic and unaccented words), 80% of the words were paroxytone, whereas approximately 17% of the words were oxytone and 3% were proparoxytone. The paroxytone pattern can thus be considered as the default accentual pattern in Spanish.

The normative spelling system establishes the following rules for placing Spanish written accent. In oxytone words, where stress is on the last syllable, a written accent is required if the word ends with -n, -s or with a vowel (e.g. *educación*, *voté*). In paroxytone words, where stress is on the penultimate syllable, a written accent is required if the word ends with a consonant other than -n and –s (e.g. *cárcel*, *carácter*). Finally, in proparoxytone words, where stress is on the antepenultimate syllable, a written accent is always required (e.g. *tránsito*).

It has been frequently noted that in Spanish orthography, the most common spelling mistake is the omission of the written accent. This comes out as the result of several investigations on different Spanish written corpora (e.g. [2], [3], [4]). In the case of Costa Rica (the variety under study in this investigation), similar results were found in a study with children (primary school) [5]. As far as adults are concerned, [6] found that 84% of the spelling errors in an adult Costa Rican written Spanish corpus (COCAE) were due to the omission of the written accent. Besides other reasons (e.g. education level, lack of attention, etc.), the difficulty in placing the written accent might be explained by the difficulty in perceiving stress. In that respect, [7] mentioned that 46% of the participants (students at the very beginning of their studies in *Pedagogy in Spanish*) had difficulties in identifying the stressed syllable. The possible relationship between spelling errors and stress perception (among other variables) was examined in 10-year children in [8]. They found that what they called the "stress awareness" (i.e. identification of the stressed syllable in non-words) was a good predictor of the spelling errors. They concluded that the prosodic skill (i.e. stress sensitivity) influenced the performance in word spelling task in 10-years children.

The aim of this investigation is to examine whether the adults' difficulty in placing the written accent is related to their ability in perceiving stress. The participant's education level (academic and non-academic) is also taken into account in this study. Moreover, as far as the stimuli are concerned, we consider the lexical status (words and non-words), the accentual pattern (proparoxytone, paroxytone and oxytone words) and the length (2, 3 and 4 syllables).

2. Method

2.1. Participants

Thirty-two Costa Rican participants (all from the San José region) took part in this experiment. They were divided into two groups. The first group was composed of 16 non-academic participants (henceforth "non-academic") (5 males and 11 females, age range: 18-70 years, mean age = 42.75). All of them had attended middle- or high-school level education, where they were taught Spanish grammar and orthography. The second group was composed of 16 students and professors from the faculty of Arts and from the faculty of Sciences of the Universidad de Costa Rica (henceforth "academic") (6 males and 10 females, age range: 19-51 years, mean age = 29.19).

2.2. Material

The auditory stimuli used in this experiment consisted of 96 items (48 real words and 48 non-words), recorded by a Costa Rican female speaker. The stimuli were prepared according to the following criteria. The real words were 48 Spanish words of two, three and four syllables, divided in proparoxytone (PP;

---

[1] Note that some adverbs ending in *-mente*, have two accented syllables (e.g. *comúnmente* [ko,mun'mente], *commonly*) and that some words, due to enclitics, present a stress on the forth to last syllable (e.g. *comiéndoselo* [ko'mjendoselo], *eating it*).

e.g. *únicos*)[1], paroxytone (P; e.g. *acabo*) or oxytone words (O; *además*)[2]. Twenty-one words required a written accent (e.g. *líder*) and 27 did not (e.g. *rosas*). The words with written accent were mainly taken from the COCAE corpus, as they generated a large number of errors ([6]). As far as the non-words were concerned, they were stimuli of two, three and four syllables (with PP, P or O accentual pattern) that were created by combining (in a random way but following the Spanish phonotactics) the syllables of the real words. This set comprised 27 non-words with written accent (e.g. *tébar*) and 21 without it (e.g. *meper*)[3].

2.3. Procedure

The experiment was run online with the *Labguistic* platform (www.labguistic.com). It was divided into two parts: *Perception* and *Spelling*. In the first part (Perception), participants heard a stimulus and had to indicate the position of stress (on the last syllable (O), the penultimate syllable (P) or the antepenultimate syllable (PP)). In the second part (Spelling), participants heard a stimulus and had to write it (with the instruction to place the written accent, if necessary).

The exact same stimuli were played in both parts (although not in the same order). Each part was composed of three sections, each one corresponding to the two, three or four syllable stimuli. Words and non-words were mixed together and presented in a different random order for each participant. The participants could hear each stimulus twice, if needed.

2.4. Data analysis

The responses given in the *Spelling* part were corrected according to the following criteria. We considered as correct responses the words (and non-words): 1) with two possible spellings, as there was no way to decide which spelling was the correct one (e.g. v/b; *boté* vs *voté*); 2) with a substitution of a letter with no implication for the presence of the written accent (e.g. *tía* instead of *día*); 3) with a substitution of a letter with an implication for the presence of the written accent (e.g. *rodes* instead of *ródez*); 4) with the addition of a letter with no implication for the presence of the written accent (e.g. *púdiro* instead of *údiro*), 5) with the omission of a letter with no implication for the presence of the written accent (e.g. *pasesa* instead of *pasesas*). We excluded the words (and non-words) with the addition or omission of a letter with an implication for the presence of the written accent (e.g. *mapovol* instead of *mapovo*; *simeracta* instead of *simeráctar*). Moreover, we collected the correct/incorrect responses of the *Perception* part.

We analyzed the data by means of mixed-effects logistic regression models in which participants and stimuli were entered as random terms ([10], [11]). For the sake of clarity, the results and figures are presented in percentages, although all statistical analyses have been performed on raw data.

---

[1] Note that disyllabic words could not be proparoxytone.
[2] We made sure that the lexeme frequency (taken from [9]) was not only similar across the three pattern ($F_{(2, 45)} = 0.93$, n.s), but also in the two, three and four syllable words ($F_{(2, 45)} = 0.88$, n.s).
[3] Although non-words do not exist in Spanish, they would follow the same Spanish accentuation rules as real words. For example, the oxytone non-word [me'peɾ] is written *meper*, and the paroxytone non-word ['teßaɾ] is written *tébar*.

3. Results and discussion

3.1. Spelling performance

A first model was run with the spelling correct/incorrect response as a dependent variable and with the following predictors: the participant's education level (academic/non-academic), the stimulus lexical status (words/non-words), the stimulus accentual pattern (PP, P, O), the stimulus length, and the participant's perception response (correct/incorrect). Given that the stimulus length had no effect and did not interact with other variables, a new analysis was run without it and without all non significant interactions.

3.1.1. Role of lexical status

As can be seen in Figure 1, results showed an effect of lexical status, with more correct responses for real words than for non-words ($F_{(1, 3033)} = 33.08$, $p < .001$), independently of the other variables. Thus, it seems easier to correctly spell (i.e. to place or not the written accent) the real Spanish words than the invented words. It is also worth it noting that the scores are very high (88.36% in average), which suggests that the task is not particularly difficult for the participants. Moreover, it is important to observe the absence of an interaction between the lexical status and the educational level, which means that the difference between words and non-words is similar in academic and non-academic participants.
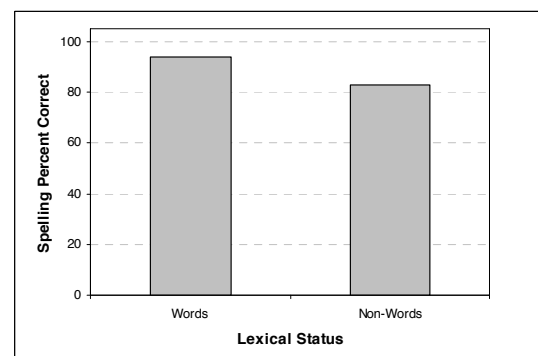

Figure 1: *Spelling percent correct as a function of lexical status.*

3.1.2. Role of education level and accentual pattern

Figure 2 presents the spelling percent correct as a function of the accentual pattern and education level. Despite the very good global performance, results show an effect of education level ($F_{(1, 3033)} = 12.66$, $p < .001$): academic participants present more correct responses (93.38%) than non-academic participants (83.34%)[4]. We also observed an effect of accentual pattern ($F_{(1, 3033)} = 4.26$, $p < .05$). Post-hoc analyses show more correct responses for PP (90.60%) and O (91.64%) than for P (85.33%) ($p < .05$).

---

[4] Given that age and gender were not similar across academic and non-academic participants, we ran a model with age and gender as predictors to make sure that these variables were not responsible for the differences observed between academic and non-academic participants. None of the variables (or their interaction) has an effect on the spelling response.

More interestingly, we note an interaction between the education level and the accentual pattern ($F(1, 3033) = 4.08$, $p < .05$). As can be seen in Figure 2, the difference between academic and non-academic is more important in PP than in P or O ($p < .05$). Moreover, academic participants present more correct responses for PP and O than for P, whereas non-academic participants present more correct responses for O than for PP or P. In other words, it seems that academic participants have more difficulties with P stimuli, while non-academic participants have more difficulties with PP and P stimuli.
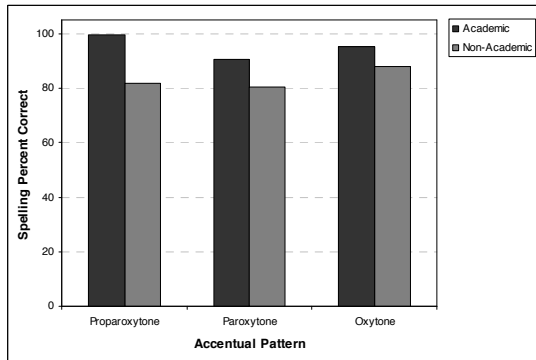

Figure 2: *Spelling percent correct as a function of accentual pattern and education level.*

3.1.3. Role of Stress perception

Interestingly, we observe an effect of the perception response on the spelling responses ($F(1, 3033) = 14.13$, $p < .001$). As can be seen in Figure 3, the stimuli with a correctly identified stress are more likely to be correctly written (i.e. with or without the written accent) than the incorrectly perceived stimuli.
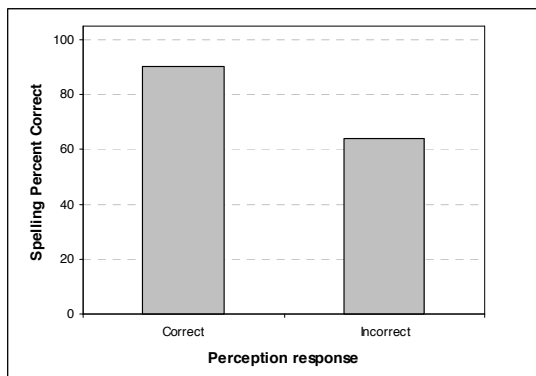

Figure 3: *Spelling percent correct as a function of the perception response.*

However, as shown in Figure 4, the effect of the perception response is modulated by the accentual pattern ($F(1, 3033) = 8.06$, $p < .001$). In fact, the effect of perception is smaller in P than in PP or O (independently of the education level) ($p < .05$). Finally, results show an interaction between the perception response and the education level ($F(1, 3033) = 17.28$, $p < .001$).
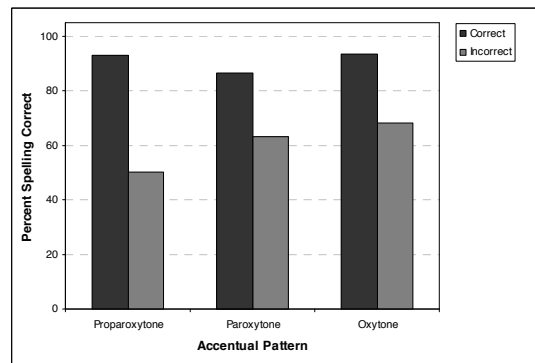

Figure 4: *Spelling percent correct as a function of the accentual pattern and the perception response.*

As can be seen in Figure 5, the difference between correct and incorrect perception responses is bigger in academic than in non-academic participants ($p < .05$). In other words, the perception of stress seems to play a more important role in academic than in non-academic participants.
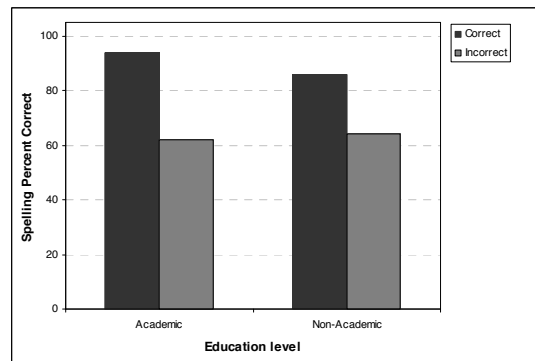

Figure 5: *Spelling percent correct as a function of the education level and the perception response.*

In order to explain the bigger importance of stress perception in academic than in non-academic participants, we ran an analysis with the perception response as the dependent variable and with the following predictors: the participant's education level (academic/non-academic), the stimulus lexical status (words/non-words), the stimulus accentual pattern (PP, P, O), the stimulus number of syllables. The results show, among other effects[1], an effect of education level: academic participants are better in identifying stress position than non-academic participants (98.63% and 88.54% respectively).

Figure 6 presents the spelling percent correct as a function of perception percent correct and education level. As can be seen, academic participants present less variability in their perception responses than non-academic participants (the standard deviation for academic participants is 1.65 and it is 15.73 for non-academic participants).

---

[1] Besides the effect of education level ($F(1, 3040) = 3.39$, $p < .001$), results of stress perception show an effect of lexical status, with more correct responses for words than for non-words ($F(1, 3040) = 5.29$, $p < .05$), an effect of the number of syllables, with more correct responses for disyllabic stimuli ($F(1, 3040) = 16.44$, $p < .001$), but no effect of accentual pattern.

Therefore, the fact that the academic participants' perception performance is not only higher, but also more invariable than the non-academic participants' performance may explain why perception plays a more important role in academic than in non-academic participants.
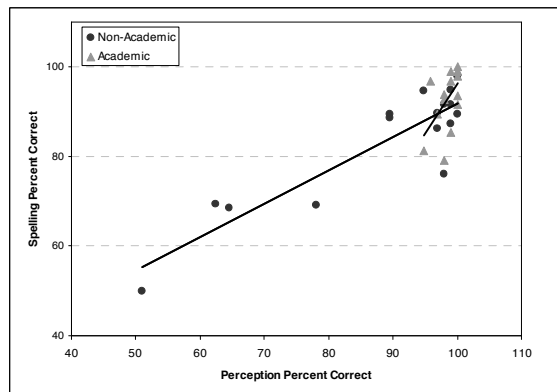

Figure 6: *Spelling percent correct as a function of perception percent correct and education level.*

3.2. Error analysis

In the error analysis, we examined whether the errors (N = 354) came from the addition of the written accent (henceafter "Added Written Accent"; e.g. *fluidéz* instead of *fluidez*) or from the absence of the written accent (henceafter "Missing Written Accent"; e.g. *carcel* instead of *cárcel*). Results show first that *Missing Written Accent* errors (N = 236) are more frequent than *Added Written Accent* errors (N = 118) (χ^2(1, N = 354) = 39.33, p < .001). In other words, participants missed more written accents than they added.

Then, an analysis was run with the error type (Missing/Added Written Accent) as dependent variable and with the following factors: education level (academic/non-academic), lexical status (words/non-words) and accentual pattern (PP, P, O) and the interactions between these variables. No main effect was significant. Only the interaction between education level and lexical status was significant (F(1, 350) = 6.40, p < .05). Figure 7 presents the percent *Added Written Accent* error as a function of lexical status and education level.
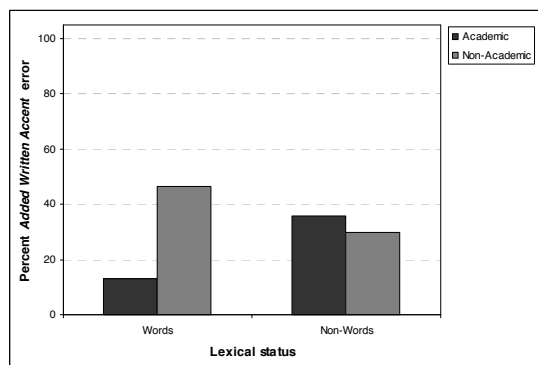

Figure 7: *Percent "Added Written Accent" error as a function of lexical status and education level.*

As can be seen, non-academic participants present more *Added Written Accent* errors than academic participants in words, while they present in non-words the same number of *Added Written Accent* errors as the academic participants. In other words, non-academic participants tend to erroneously add a written accent to a greater extent than academic participants, especially in real words.

4. General discussion

The aim of the investigation was to examine whether the adults' difficulty in placing the written accent in Spanish words was related to their ability in perceiving stress. We also took into account the following variables: the participant's education level, the stimulus lexical status, accentual pattern and length.

While results showed no effect of the stimulus length, they showed an effect of lexical status (i.e. more correct spellings for words than for non-words). This finding might suggest that the presence of the written accent is stored in the mental representation of the words. However, further research is still needed to confirm this hypothesis. An effect of education level was also observed, which indicates that academic participants mastered to a greater extent the Spanish written accentuation rules than non-academic participants. More interestingly, we found an effect of accentual pattern, with less correct responses for paroxytone stimuli than for proparoxytone or oxytone stimuli. We already mentioned that the paroxytone pattern is the default accentual pattern in Spanish. Yet, determining how frequently the paroxytone words require a written accent may explain why this pattern presents more errors. To do this, we analyzed a corpus composed of formal written Spanish texts (CODIMEP-CR-XXI). It contained 124'000 tokens, which corresponded to 11'370 types (i.e. different words). We found that 16% of the words have a written accent. Among those, 53% were oxytone, 31% were proparoxytone and 16% were paroxytone. It is important to note that, among the paroxytone words with written accent, only 15% corresponded to words ending in consonants other than -n and -s. (e.g. *dólar*), while 85% corresponded to words ending with hiatus (e.g. *día*). Yet, our paroxytone stimuli with written accent always showed this infrequent spelling (e.g. *líder*, *cárcel*, *fácil*, etc.). Moreover, taking into consideration that the default pattern in Spanish is paroxytone [1] and that only 16% of Spanish words require written accent, from which 85% are proparoxytone or oxytone (corpus CODIMEP-CR-XXI), we can assume that Spanish paroxytone words with written accent are much less frequent that paroxytone words without written accent. Yet, our stimuli included the same number of paroxytone items with and without written accent (words and non-words). This might explain why there are more errors in the spelling of paroxytone stimuli.

Finally, we observed an effect of the stress perception in the assignment of a written accent: stimuli with a correctly identified stress are more likely to be correctly written (i.e. with or without written accent) than the incorrectly perceived stimuli. However, this effect is weaker in paroxytone words and in non-academic participants. Taken together, these results confirm the hypothesis that the adults' difficulty in placing the written accent in Spanish is related to their ability in perceiving stress. In other words, there is a relationship between prosodic and written skills, as it was showed for reading in [8].

5. References

[1] Quilis, A., "Tratado de fonética y fonología españolas", 2nd. ed., Gredos, Madrid, 1999.

[2] Mesanza, J., "Palabras que peor escriben los alumnos (inventario caográfico)", Escuela Española, Madrid, 1990.

[3] Pujol, M., "Análisis de errores grafemáticos en textos libres de estudiantes de enseñanzas medias". Doctoral dissertation, Departament de Didàctica de la Llengua i la Literatura, Universitat de Barcelona, 1999. Online: http://hdl.handle.net/2445/41392, accessed on 4 Dic 2013.

[4] Pujol, M., "La ortografía", in: S. Torner and M. P. Battaner [Eds.], El corpus PAAU 1992: estudios descriptivos, textos y vocabularios, 29-65, Universitat Pompeu Fabra, Barcelona, 2005.

[5] Murillo Rojas, M., "Vocabulario cacográfico. Pautas para la enseñanza de la ortografía en la escuela primaria costarricense". Káñina, Revista de Artes y Letras de la Universidad de Costa Rica, 30(1):59-70, 2006.

[6] Jara Murillo, C. V., "COCAE: Corpus Cacográfico Adulto del Español de Costa Rica", Research Report, Project No. 745-B2-A13, Universidad de Costa Rica, 2013. Online: http://www.kerwa.ucr.ac.cr/handle/10669/8928, accessed on 4 Dec 2013.

[7] Henry, E., "Dificultades en la percepción del acento", Revista Lingüística Teórica y Aplicada 21, Concepción, Chile, 1983.

[8] Defior, S., Gutiérrez-Palma, N. and Cano-Marín, M. J., "Prosodic awareness skills and literacy acquisition", Journal of Psycholingusitic Research, 41: 285-294, 2012.

[9] Alameda, J. R. and Cuetos, F., "Diccionario de frecuencias de las unidades lingüísticas del castellano". Oviedo: Servicio de Publicaciones de la Universidad de Oviedo, 1995.

[10] 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.

[11] Bates, D. M. and Sarkar, D., "lme4: Linear mixed-effects models using S4 classes", R package version 2.6., www.r-project.org, 2007.