Degrees of Phonetic Similarity and Analogically-Driven Stress Perception in Spanish

Timothy L. Face
University of Minnesota, USA

ABSTRACT

This paper examines the influence of a similar real Spanish word on the perception of stress in a nonsense word, focusing on the degree of phonetic similarity between the two. Nonsense words were created and synthesized that, while similar to a real word, would be predicted to have a different stress pattern from the real word if the real word does not have an analogical effect on stress perception. The results show varying degrees of influence of the real word on the perception of stress in the nonsense words, with the degree of influence depending on the phonetic similarity of nonsense word to the real word. The results add to the understanding of the ways in which the structures of Spanish are cues to stress perception, and also lend support to recent claims that whole words stored in the lexicon are used actively through analogy by speaker/hearers.

1. INTRODUCTION

In studying the nature of Spanish stress, experimental studies have focused principally on the acoustic correlates of stress (i.e. fundamental frequency, duration, and intensity) both in production studies (e.g. [1] and [2]) and in perception studies (e.g. [3] and [4]). This is only one aspect of a larger picture, however. In [1], Quilis states that "the sensation of a physical stimulus is channeled through the structures of a language" (translation mine). In [5] I conducted an experiment to test whether syllable structure (specifically syllable weight) affected the perception of stress in nonsense words where a synthesized voice was controlled so as to make all syllables of equal acoustic prominence. The results indicated that different syllable weight combinations lead to a consistent agreement between native Spanish-speaking hearers as to which syllable is stressed. The present study extends my previous study by considering another way in which the structure of Spanish leads to the perception of a particular syllable as stressed.

2. EXPERIMENTAL METHODS

In order to test whether the stress pattern of a similar word has an effect on the perception of Spanish stress, ten real trisyllabic Spanish words ending in a consonant and having final stress (e.g. *hospital* 'hospital') were chosen. These were the similar words that could have an influence on the perception of stress in a segmentally similar nonsense word. In order to create nonsense words similar to the real words but which should have a different stress pattern if there were no influence of the similar word, nonsense words were synthesized which were segmentally identical to the real words except that they lacked the final consonant of the real word (e.g. [ospita] from *hospital* 'hospital'). Since the nonsense words end in a vowel, they are predicted to be perceived as having penultimate stress if the similar real word does not exert an analogical influence on stress perception. Each word was synthesized twice using the MBROLI speech synthesizer ([6]). In one case, the now final vowel was synthesized as a word final vowel. In the other case, the now final vowel maintained the acoustic characteristics of the vowel preceding the final consonant of the real word. This allowed for different degrees of phonetic similarity of the nonsense word to the real word.

While eliminating the final consonant of a real word with final stress sets up a clear prediction of what the results should look like if real words do have an influence on stress perception and if they do not, there is something else to consider. Since the nonsense words are segmentally identical to real words except for the lack of the final consonant, the processing of the nonsense word matches the real word up until the nonsense word ends, which should clearly activate the real word for the subject (see [7] and [8], as well as Section 4, for information on word recognition and activation). In order to test whether a similar word could have an influence on stress perception without the possibility of being processed as identical to the real word, a second set of nonsense words was created. The second set of nonsense words was identical to the first set except that a final [s] was added to the end of each (e.g. [ospitas] from *hospital* 'hospital'). Since words ending in [s] generally have penultimate stress, there should again be a difference between perceived penultimate stress if the real word is not an influence and perceived final stress if the real word is an influence. The predictions, then, are the same as with the vowel-final nonsense words

The thirty synthesized nonsense words (ten vowel-final words synthesized with the final vowel having the qualities of a word-final vowel, ten vowel-final words where the final vowel has the qualities of the corresponding vowel in the similar real word, and ten [s]-final words) were interspersed randomly with other nonsense words lacking acoustic stress as well as with a control group of words that did have an acoustically stressed syllable. This control group was included to be sure that when an acoustically stressed syllable was present, the subjects could accurately identify it. Any potential subject who did not accurately identify the acoustically stressed syllable at least 90% of

ISBN 1-876346-48-5

the time was excluded from the study. In this way the perception data obtained would be from speaker/hearers who can accurately perceive and identify a stressed syllable. If a speaker/hearer were unable to identify the stressed syllable, then her judgment of stress in the test words would not provide meaningful and relevant data.

The subjects who participated in the experiment were ten native Spanish-speaking graduate students at the University of Minnesota who had not lived in the United States before attending graduate school. All grew up in a monolingual environment, although they came from various parts of the Spanish-speaking world. While the same three acoustic correlates of stress (i.e. fundamental frequency, duration, and intensity) are used in different dialects of Spanish, there may be dialectal differences in the importance of each. In the present study, however, this is not a factor since the experiment neutralizes these acoustic correlates of stress in order to test the role of non-acoustic factors on stress perception. The issues addressed in the present study apply equally to all varieties of Spanish. All of the subjects were naïve with respect to the purpose of the study and had no training in Spanish phonology.

3. RESULTS

Recall that without an analogical influence of the similar real word, all target nonsense words would be predicted to be perceived as having penultimate stress. If, on the other hand, the similar real word does exert an analogical influence on stress perception, then the target nonsense words would be predicted to be perceived as having final stress. The number of perceived cases of initial, penultimate, and final stress are shown in Table 1 for each of the three sets of target nonsense words.

Perceived stressed syllable	Vowel-final (quality of final vowel in real word)	Vowel-final (quality of a word-final vowel)	s-final
Initial	12	12	18
Penultimate	16	29	45
Final	72	59	37

Table 1: Perceived stressed syllables by set of target nonsense words.

The data in Table 1 show that for both categories of vowel-final words, final stress is perceived more often than penultimate stress and initial stress combined. This indicates an analogical influence of the similar real word on the perception of stress in the target nonsense words, since without such an influence a vowel-final word would be predicted to have perceived penultimate stress. Clearly, however, there is a substantially higher rate of perception of final stress in those vowel-final words where the final vowel has the quality of the vowel in the real word than in those cases where the vowel has the quality of a word-final vowel. This seems to indicate that the degree of phonetic similarity is a factor in stress perception. When the final vowel of the target nonsense word has the quality of the corresponding vowel in the real word (which is pre-consonantal rather than word-final), the nonsense word is more similar to the real word than when the final vowel of the nonsense word has the quality of a word-final vowel.

In the case of the [s]-final nonsense words, penultimate stress is perceived more often than final stress. Since penultimate stress is predicted for an [s]-final word if there is no analogical influence of the similar real word, it can be safely assumed that the similar real word has less of an effect on the [s]-final words than on the vowel-final words. Even so, for the [s]-final words, final stress is perceived in 37 of 100 cases. This is a higher rate of perception of final stress than would be expected if the similar real word had no effect on the perception of stress in the [s]-final words. In fact, in [5] I reported for an earlier study that for trisyllabic vowel-final nonsense words which were not similar to a real word and which lacked an acoustically stressed syllable, final stress was perceived only 13.5% of the time, with penultimate stress being perceived 76% of the time.

While in all three categories of target nonsense words there seems to be an analogical influence of the similar real word on stress perception, it is also undeniable that the degree of phonetic similarity of the nonsense word to the similar real word is a factor. A Chi-Square test applied to Table 1 shows that the difference in perceived stress between categories of target nonsense words is highly significant ($p<.0001$). The more similar the target nonsense word is to the similar real word, the more likely that final stress will be perceived. In this way the perceived stress pattern (to more or less of a degree, depending on the category of the target nonsense word) follows that of the similar real word with final stress and breaks with the pattern of penultimate stress for vowel-final and [s]-final words.

4. DISCUSSION

The experimental results in the preceding section show that a similar real word has an analogical influence on the perception of stress in a nonsense word. Furthermore, the degree of phonetic similarity of the nonsense word to the real word determines the strength of the analogical influence. The more similar the nonsense word is to the real word, the more likely that the stress pattern of the real word will be perceived in the nonsense word. But what are the larger implications of these experimental results for the linguistic system of the speaker/hearer?

First, the results fit nicely with the Neighborhood Activation Model of spoken word perception proposed in [7]. According to the Neighborhood Activation Model, as a hearer hears the beginning of a spoken word, it activates all words that begin the same way. For example, if a speaker/hearer of Spanish hears the initial sequence [ka], it activates the words *casa* ‘house’, *caso* ‘case’, *cama* ‘bed’, cana ‘cane’, carta ‘letter’, cada ‘each’, cabo ‘cape’, etc. As

ISBN 1-876346-48-5

the hearer processes more of the spoken word, the words that still begin the same as the spoken word remain activated, while those that do not begin the same way do not remain activated. So, following the Spanish example, if after [ka] the next segment heard is [s], only words that begin with [kas] will remain activated (e.g. *casa* 'house', *caso* 'case'). This process continues until only one word remains activated, and that is the word perceived by the hearer. In the present study, subjects were presented with nonsense words that begin like a real word. By the time they hear the end of a vowel-final nonsense word, only one real word remains activated. While the final segment of the real word is not part of the nonsense word, neither has any segment differing from the real word been perceived. Therefore the real word remains activated. Because of this, the lone real word that remains activated is likely to exert an influence on the perception of the nonsense word. This is why speakers quite frequently perceive the stress pattern of the real word rather than that predicted for vowel-final words.

The fact that the nonsense words in which the final vowel had the same quality of the corresponding vowel in the real word more often were perceived as having the stress pattern of the real word than those nonsense words in which the final vowel had the quality of a word final vowel is also easily explainable. While segmentally the two vowel-final nonsense words leave the same single real word activated, there is a phonetic difference in the quality of the final vowel between the two nonsense words, and hearers may be attuned to this phonetic distinction. Therefore, while the lone activated real word may influence stress perception in both cases (since it is activated for both vowel-final nonsense words), the fact that the vowel of the real word is phonetically more similar to the corresponding vowel of one of the nonsense words than to that of the other means that the real word is more likely to influence stress perception of the word with the phonetically most similar vowel. In other words, in terms of phonological segments the two nonsense words are identical and should be equally susceptible to the influence of the activated real word, but the fact that there is a phonetic difference between the nonsense words leads to a different degree (or different likelihood) of influence from the real word.

In the case of the [s]-final nonsense words, upon perception of the vowel preceding the final [s] there is one real word activated. Once the [s] is perceived, no real words remain activated that match the speech stream. Therefore it is far less likely for any real word to exert an analogical influence on the perception of the nonsense word. Nonetheless, final stress is perceived 37% of the time in the [s]-final nonsense words. This indicates that the real word that remains activated the longest may still influence the perception of the nonsense word, though it is far less likely than in cases where the real word remains activated through the completion of perception of the nonsense word. So while activated words are highly likely to serve as analogical models for perception, a previously activated word may also serve as an analogical model, though the likelihood is not great.

While is it possible to place the results of the experimental study of the present study in terms of a model of spoken word recognition, it is also possible to place them in terms of a larger model of lexical storage and structure. The results fit particularly well within Bybee's usage-based model ([9]). Within this model, the lexicon stores tokens of language use in full phonetic detail. Individual lexical entries are linked to others by their phonetic/phonological form, their semantic/pragmatic meaning, etc. Each time a speaker/hearer experiences language, this is stored in the lexicon. The structure of the lexicon is the complex network of links between items similar to each other in one way or another (e.g. form or meaning). The links that are formed lead to analogical influence of one form on the other, as Bybee discusses and as has been shown by various experimental studies. But how can a usage-based model deal with the experimental results reported in Section 3?

When considering the usage-based model, it is form (rather than meaning) that is of interest in the present discussion. The nonsense words in the present study are similar in form to real words in the language. These (and all other real words that speaker/hearers have experienced) are stored in the lexicon. The lexicon, then, represents a speaker/hearer's experience with her language. Therefore, the existing sound patterns that are found in the lexicon are the knowledge that the speaker/hearer has of the sounds of her language. So, similarly to the Neighborhood Activation Model of spoken word recognition, when a word is heard and processed, the similar words in the lexicon are those that are most likely to be perceived. Rather than the activation of particular words, numerous similar words (even if not segmentally identical up through the portion of the word processed at any given point) will be evoked by hearing a word. These may be other words that have a similar form, a similar meaning, etc. Thus, while the model may be less specific about exactly how the word recognition process occurs than is the Neighborhood Activation Model, it is able to account for more than just the recognition of a single spoken word. It also accounts for issues such as semantic and morphological priming.

With respect to the present study, the usage-based model accounts for the results of the perception experiment because the target nonsense words are similar to real words. Since the word being heard by the hearer is then linked to the real word by its form, the real word can serve as an analogical model for perception. The closer the form, the more likely that there will be an analogical effect. In the case of the [s]-final words in the present study, they are still related to the real words due to their nearly identical form (recall that only the final consonant is different). Therefore there is a natural explanation for an analogical effect in these cases. Of course, the form differs from that of the real word at the end of the nonsense word, and this divergence from the real word stored in the lexicon means that there is a lesser chance of an analogical influence. The usage-based model, then, accounts easily for different degrees of influence based on the degree of similarity of form. This is an advantage over the Neighborhood Activation Model, in which a word is activated or not and does not easily

 ISBN 1-876346-48-5

incorporate differing degrees of activation. In addition, in the case of a nonsense word that is quite different from any real word, the usage-based model still accounts for the stress pattern perceived because the nonsense word will still be related in form to other words (e.g. in that it is vowel-final), where the Neighborhood Activation Model is only able to deal adequately with the recognition of real (or known) words.

Another benefit of the usage-based model that merits mention here is that it goes beyond speech perception. The tokens of language use stored in the lexicon are also responsible for speech production. Especially when there is variation, the patterns of variation found in the language are mirrored by the variation found in production of nonsense words. A prime example is Aske's study of Spanish stress placement in nonsense words ([10]). Aske shows that when stress placement is variable based on the segmental content of a word, the same variability (or lack thereof) is found in the placement of stress in nonsense words by Spanish speakers, based on the segmental content of the nonsense words. So in non-verb nonsense words ending in [n], when the preceding vowel is [e], Spanish speakers have a fairly even distribution between penultimate stress and final stress. If the preceding vowel is any vowel other than [e], stress placement is overwhelmingly final. This mirrors the patterns of stress placement found for real Spanish non-verbs ending in [n]. So the usage-based model not only accounts for stress perception, but also for how speakers determine stress placement in production, even when it is variable.

5. CONCLUSIONS

The experimental results presented in the present study show that stress perception in nonsense words is influenced by the stress pattern of a similar Spanish real word. Furthermore, the degree of phonetic similarity between the real word and the nonsense word plays a role. The more similar the nonsense word is to the real word, the more likely the real word is to affect stress perception in the nonsense word. These results were considered in terms of the Neighborhood Activation Model of spoken word recognition (as presented in [7]) and in terms of a usage-based model of phonology (as presented in [9]) that relies heavily on the storage of tokens of language experience in the lexicon in full phonetic detail. While the experimental data can be incorporated into both models, the usage-based model is more dynamic and is better equipped to deal with the varying degrees of analogical influence from the real word on the perception of the nonsense word. Furthermore, the usage-based model accounts for related production data, such as those presented in [10].

While other studies have used Spanish data to argue for a usage-based model of phonology, this is the first to use perception data. It also extends my previous study ([5]) that examined the role of syllable weight in the perception of Spanish stress. Further studies should continue this line of research investigating the non-acoustic factors that lead speaker/hearers of Spanish to perceive stress on a particular syllable. While segmental issues (such as those construed in [5] as being instantiations of differing syllable weights) and degree of similarity to another word (or, in the case of nonsense words, to a real word) play a role in Spanish stress perception, there are likely other non-acoustic factors as well that enter into the cognitive processes of a Spanish speaker/hearer and influence stress perception.

REFERENCES

[1] A. Quilis, "Caracterización fonética del acento español," *Travaux de Linguistique et de Littérature*, vol. 9, pp. 53-72, 1971.

[2] A. Quilis, *Tratado de fonología y fonética españolas*, Madrid: Gredos, 1993.

[3] E.V. Enríquez, C. Casado, and A. Santos, "La percepción del acento en español," *Lingüística española actual*, vol. 11, pp. 241-269, 1989.

[4] J. Llisterri, M.J. Machuca, C. de la Mota, M. Riera, and A. Ríos, "Algunas cuestiones en torno al desplazamiento acentual en español," Paper presented at La tonía: dimensiones fonológicas y fonéticas, Mexico City: Colegio de México, 2002.

[5] T.L. Face, "The role of syllable weight in the perception of Spanish stress," in *Hispanic Liguistics at the Turn of the Millennium*, H. Campos, E. Herburger, A. Morales-Front, and T.J. Walsh, Eds., pp. 1-13. Somerville, MA: Cascadilla Press, 2000.

[6] T. Dutoit, V. Pagel, N. Pierret, F. Bataille, and O. Van der Vrecken, "The MBROLA project: toward a set of high-quality speech synthesizers free of use for non-commercial purposes," *Proceedings of ICSLP'96*, pp. 1393-1396, 1996.

[7] P.A. Luce and D.B. Pisoni, "Recognizing spoken words: the neighborhood activation model," *Ear & Hearing*, vol. 19, pp. 1-36, 1998.

[8] D.B. Pisoni, H.C. Nusbaum, P.A. Luce, and L.M. Slowiaczek, "Speech perception, word recognition and the structure of the lexicon," *Speech Communication*, vol. 4, pp. 75-95, 1985.

[9] J. Bybee, *Phonology and Language Use*, Cambridge: Cambridge University Press, 2001.

[10] J. Aske, "Disembodied rules versus patterns in the lexicon: testing the psychological reality of Spanish stress rules," in *Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistics Society*, K. Hall, J-P. Koenig, M. Meacham, S. Reinman, and L.A. Sutton, Eds., pp. 30-45. Berkeley: Berkeley Linguistics Society, 1990.