From Tone to Accent: the Tonal Transfer Strategy for Chinese L2 learners of Spanish

Yudong Chen

University of Illinois at Urbana Champaign
ychen28@uiuc.edu

ABSTRACT

The transfer of tone and intonation in second language acquisition is a relatively unexplored field. In this paper, I investigate the acquisition of Spanish prosodic patterns by native speakers of Standard Mandarin Chinese. Pitch plays very different linguistic roles in Mandarin Chinese and in Spanish. Chinese is a language in which the tonal contour of individual syllables is lexically contrastive. Spanish, on the other hand, is an intonational language, where tonal contours characterize utterances and convey pragmatic functions. Conversely, Spanish has lexically contrastive stress which serves as anchoring points for local pitch excursions or pitch accents. The same word will appear with different tonal contours depending on its position in the phrase and the pragmatics of the discourse. In particular, in words in prenuclear position in a declarative sentence the stressed syllable is often characterized by a rise which starts in the stressed syllable and continues in the following syllable [2], [3].

In this paper we find strong evidence for the hypothesis that Mandarin speakers learning Spanish as a second language interpret the contours of Spanish words in citation form as a lexical property of individual syllables. For instance, a word like *sa**li**da*, with stress on the penultimate syllable, will be interpreted as having a rising tone on the penultimate syllable (= tone 2 in Chinese) and a falling tone on the final syllable (= tone 4). This interpretation of intonation as lexical tone leads these learners to employ contours with a tonal rise in the stressed syllable and a fall on the post-tonic also in prenuclear position. This tonal transfer strategy from L1 to L2 is shown as a robust feature.

Keywords: Tone, Transfer, Stress, Spanish, L2

1. INTRODUCTION

Speakers of Standard Mandarin (SM) usually have no intuition of the position of stress of SM. Acoustic phonetic studies of stress in SM do not offer consistent results. According to Yan and Lin [4], when SM words are read in isolation, the last syllable has the longest duration. However, Wang and Wang [5] show that when Mandarin words are read in a carrier sentence, the first syllable is usually longer than other syllables. On the other hand, stress is a contrastive feature at the lexical level in Spanish. Stressed syllables act as anchoring points for intonational pitch accents, but the specific pitch accent associated with a given word depends on pragmatic factors and the position of the word in the phrase. Specifically, the delaying of F0 peaks of stressed syllables in words in prenuclear position in declarative sentences has been noticed for a long time [3].

Since speakers of SM generally have no intuition of stress, the production of the stress of a language such as Spanish has to be explicitly taught to Chinese L2 learners. SM being a tonal language, speakers of SM are sensitive to pitch changes. The claim defended in this paper is that Chinese speakers apply syllable-level F0 contours as the primary cue of Spanish stress. Two hypotheses are proposed based on this argument:

Hypothesis 1: Chinese speakers tend to relate the stress of Spanish more tightly with the F0 contour of the stressed syllable, and hence reduce the delaying of F0 peak in stressed syllables.

Hypothesis 2: Chinese speakers tend to treat pitch as the primary cue of word-stress and associate the stressed syllable with a rising tone (=Mandarin tone 2) and the immediately following syllable with a falling tone (=tone 4).

2. PAGE LAYOUT AND STYLE

2.1. Participants

40 native speakers of SM participated in this study. The participants are current undergraduate students of Beijing Foreign Studies University (BFSU) majoring in Spanish language. The participants are divided into four equal-sized groups according to

the total years of instruction received in Spanish. The participants' age ranges from 18 to 23 and are approximately balanced in gender. All participants had little or no exposure to Spanish before college.

2.2. Procedure

The experiment consisted of reading single words, sentences and a free conversation in Spanish. The data reported in this paper are collected from the sentence-reading task. The participants' speech was recorded onto a laptop computer with a microphone. The recording for each participant lasted for about 20 minutes. The sound files were analyzed with Praat to extract pitch values. The data were processed with Python. The statistical analysis and graphs were produced with R.

2.3. Material

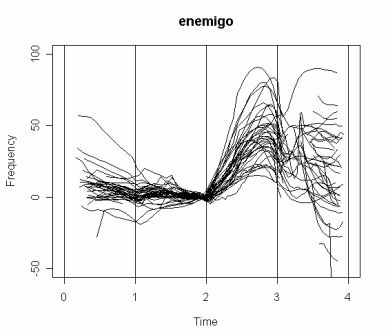
Each participant read 20 sentences, which are responses to 20 questions with broad focus. The target words are nouns in sentence-medial position and embedded within a noun phrase. The syntactic context was always either Determiner Noun (target word) Adjective or Determiner Noun (target word) Prepositional Phrase, which are neutral word orders in Spanish. This is a context where peak displacement to the post- tonic syllable is expected in Peninsular Spanish. Half of the target words have stress on the penultimate syllable and half on the antepenultimate syllable.

3. RESULT AND ANALYSIS

3.1. Pitch Contour Analysis

In Figure 1, the pitch contours of the same word (*ene**mi**go*) produced by all participants are plotted.

Figure 1: Normalized pitch contour of the word *enemigo* for all 40 participants. (Stressed syllable is in boldface).



The starting points of the pitch contour of the stressed syllable in Figure 1 are equated and set to 0 for the purpose of comparison. The vertical lines in the figures are syllable boundaries. The duration of all syllables are normalized and equated to 1. An alternative methods is to use the average syllable length as length of each syllable of the graph, however this would make the comparison across different words sharing the same stress pattern more difficult.

Figure 2: Smoothed pitch contour of the word *enemigo* for all 40 participants. (Stressed syllable is in boldface).

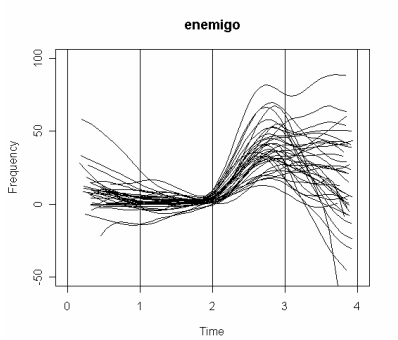


Figure 2 shows the smoothed curve corresponding to the raw pitch contours in Figure 1. The smoothing algorithm is the sm.regression proposed by Bowman and Azzalini [1], which has a normal kernel and the standard deviation is set as 1.2. The smoothed curve removes all the jittering in the original pitch contour while preserving the general pattern of the pitch contours.

The same analysis was carried out with all 20 target words. From these graphs, several observations can be made:

1. The peak of the pitch contour falls exactly on or very close to the syllable boundary between stressed syllable and post-tonic syllable.
2. The stressed syllable seems to be associated with a rising pitch contour. When the post-tonic syllable is not the last syllable of the word, it seems to be associated with a falling pitch.
3. There is no clear pitch pattern associated with the last syllable of the word.

These findings are consistent across all speakers and stress patterns.

3.2. Boundary Distance Analysis

To test the first hypothesis, the distance between the peak of the F0 contour and the syllable boundary of the stressed and post-tonic syllable was measured. Since this measurement is based on the smoothed curve instead of the original pitch contour (in order to eliminate the effects of

jittering), it is relatively easy to detect the local maximum closest to the syllable boundary. Due to the fact that the smoothed curve may introduce more variability to the location of local maximum, a higher confidence level is required for the statistical analysis (t-value bigger than 3).

An algorithm implemented in R was used to automatically detect the local maximum. The first local maximum encountered in the stressed syllable or the post-tonic syllable was considered as the peak. If no peak was detected, the sample was discarded. The average value of the peak-boundary distance was calculated for each word. The results are listed in Table 1:

Table 1: Boundary-Peak Distance of the target words

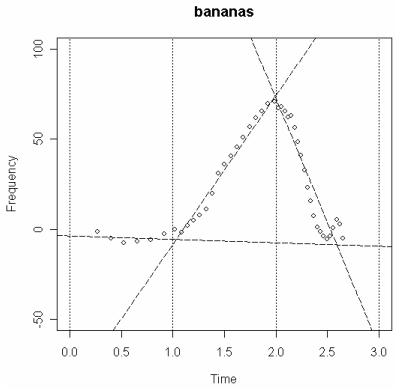
	Mean	SD	T value	DF
aguilas	-0.28	0.03	-10.12	30
camara	-0.50	0.08	-6.67	27
idolo	-0.46	0.04	-11.47	31
minimo	-0.03	0.06	**-0.53**	34
nomadas	-0.07	0.04	**-1.88**	37
numero	0.19	0.04	5.20	35
sabado	0.03	0.12	**0.25**	16
silabas	-0.20	0.07	**-2.97**	15
ba**na**nas	-0.02	0.05	**-0.45**	26
ge**me**lo	-0.04	0.05	**-0.81**	28
he**la**do	0.03	0.05	**0.58**	28
lla**ma**da	0.15	0.06	**2.67**	30
lle**ga**da	0.19	0.06	3.08	27
ma**no**lo	0.14	0.07	**2.13**	27
sa**li**da	0.05	0.07	**0.83**	27
se**ma**na	0.37	0.05	7.38	18
ene**mi**go	-0.06	0.05	**-1.07**	34
fe**no**meno	0.10	0.05	**1.76**	37
mone**de**ro	0.06	0.07	**0.76**	28
mo**no**logo	-0.16	0.03	-5.38	35
all	-0.03	0.36	**-2.08**	609

T-values larger than 3 are highlighted. Overall, the peak appears to fall slightly before the syllable boundary, but the corresponding t-value (-2.08) is smaller than the preset significance level (3). Examining each word individually, 14 out of 20 words have relatively small peak-boundary distances, which are considered not significantly different from perfect alignment between peak and syllable boundary. A more careful examination of different types of stress pattern shows that when the first syllable of the word is the stressed syllable, half of the peak-boundary distances are long enough to be considered significant; while only 2 out of the 12 words where the stressed syllable is preceded by another syllable have a significant peak-boundary distance. Based on this observation, I conclude that the first hypothesis is partially supported by the results of the experiment: the peak will always lie on or close to the end-boundary of the stressed syllable. Additional data would be needed to explain the mixed pattern of words with word-initial stress.

3.3. Stress Pattern Analysis

Based on observation of the data, I conclude that the stressed syllable is characterized by a rising pitch contour and the post-tonic syllable by a falling pitch contour. To further analyze the degree of rising and falling, a simple linear regression was conducted on each syllable taking time as independent variable and pitch value as the dependent variable. Figure 3 shows an example of an original pitch contour and a fitted line:

Figure 3: Example of original pitch contour (dotted line) and the fitted line (dash line).



The intercept and the slope of each syllable are calculated from the regression. The slopes convey information regarding the degree of rising and falling of the pitch contour in general. A positive slope means an upward movement of the pitch contour, while a negative slope means a downward movement. A first order simple linear regression is sufficient if it is assumed that each syllable is characterized by a rising, falling or flat pitch movement. A certain degree of variation may be introduced by the position of the peak if it does not fall exactly on the syllable boundary. However, from Table 1 we can see that such effect should be small. To further eliminate this effect, we require a higher confidence level (t value = 3).

The slopes of all the syllables of the words were collected and a histogram was drawn for each

syllable of all target words. An example is shown in Figure 4.

Figure 4: Histogram of slopes of pitch for each syllable of the word *fenómeno* for all the participants

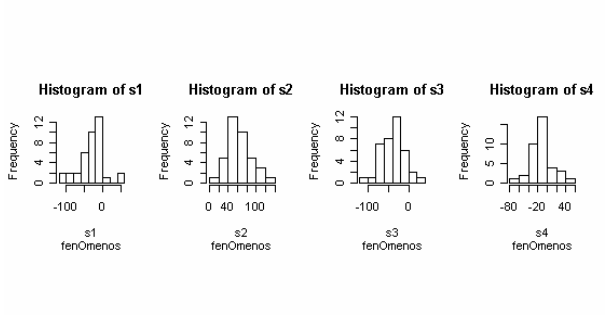


Figure 4 shows that the stressed syllable of the word *fe**no**meno* has a distribution centered on 40, which is a relatively strong rising pattern, while the post-tonic syllable has a distribution centered around -50. The distribution of the slopes in these two syllables is fairly normal.

The average value of the slopes of all the syllables of the target words is listed in Table 2. The third column corresponds to the stressed syllable. T-values larger than 3 are considered significant and are highlighted.

Table 2: Pitch Slope of all the syllables of the target word

	Mean	T	Mean	T	Mean	T	Mean	T	Mean	T
enemIgo	-9.7	**-3.5**	-2.2	-1.6	45.8	**5.3**	-21.6	-2.5		
monedEro	-12.1	**-4.6**	-6.7	**-4.4**	56.0	**5.7**	-19.3	-2.4		
banAna			-7.3	**-3.3**	55.1	**5.3**	-31.4	-2.8		
gemElo			-26.6	-1.7	54.8	**5.2**	-25.6	-2.5		
helAdo			-3.8	-1.2	48.8	**5.4**	-31.0	-2.5		
llegAda			-12.1	**-3.8**	34.1	**4.9**	-0.6	-0.1		
manOlo			-10.0	**-3.3**	59.9	**5.4**	-5.0	-0.9		
sallda			-39.7	**-4.9**	51.1	**5.4**	0.1	0.0		
semAna			-52.2	**-4.8**	47.1	**5.3**	11.5	1.9		
llamAda			-25.4	**-4.1**	42.6	**5.3**	-13.8	-1.9		
fenOmenos			-31.8	**-4.3**	64.2	**5.7**	-36.8	**-4.9**	-10.2	-2.5
monOlogo			-7.7	**-3.5**	66.3	**5.6**	-51.8	**-5.2**	-18.0	**-4.1**
Aguila					4.3	1.1	-27.0	**-4.7**	-3.7	-1.0
cAmara					0.1	0.1	-43.1	**-4.3**	-20.1	**-4.5**
mInimo					54.4	**5.0**	-28.4	**-4.4**	-30.0	**-3.7**
Idolo					0.6	0.1	-37.7	**-5.1**	3.5	0.4
nOmadas					49.2	**5.2**	-57.5	**-5.4**	-9.4	-2.8
nUmero					68.3	**5.7**	-24.6	**-3.9**	-11.5	-2.7
sAbado					-22.2	**-3.3**	-21.2	**-3.9**	-1.3	-0.4
sIlabas					-9.3	-1.3	-49.5	**-5.3**	-16.5	-2.1

The slopes are all significantly positive in the stressed syllable when the syllable is preceded by another unstressed syllable. This implies that speakers tend to associate the rising pitch contour with the stressed syllable. If the stressed syllable falls on the first syllable of the word, the pattern is less clear. The slopes are all significantly negative in the post-tonic syllable when it is followed by another unstressed syllable. This implies that speakers tend to associate a falling pitch contour with the post-tonic syllable. The last syllable of the word does not show a strong pattern, even if it is the post-tonic syllable.

4. CONCLUSION

The data from the experiment appear to support the two hypotheses proposed in section 1. Chinese speakers learning Spanish seem to interpret Spanish stress as a combination of a rising and a falling tone. The fact that the boundary of the rising and the following tone falls exactly on the syllable boundary suggests that the Chinese learners tends to associate the rising and falling pitch contour with two separate syllables. This is very different from the tonal patterns employed by native Spanish speakers, especially in words in prenuclear position. This may be considered as evidence of interpretation of Spanish intonational pitch accents as lexical syllable-level tones, associated with the stressed and post-tonic syllable of words.

The large standard deviation and small t-value of slopes associated with the final syllable may be due to the influence of the linguistic structure after the target word or may be due to the intonation of the phrase or sentence in general. The mixed pattern associated with the stressed syllable at the beginning of a word may be due to the relative short duration of word-initial syllables.

5. REFERENCES

[1] Bowman, A.W., Azzalini, A. 1997. *Applied Smoothing Techniques for Data Analysis: the Kernel Approach with S-Plus Illustrations.* Oxford: Oxford University Press.

[2] Hualde, J. I. 2002. Intonation in Spanish and the Other Ibero-Romance Languages: Overview and Status Quaestionis. *Romance Phonology and Variation*. *Selected papers from the 30th Linguistic Symposium on Romance Languages.* Amsterdam and Philadelphia, 101-115.

[3] Navarro Tomás, T. 1*944. Manual de entonación española*. New York: Hispanic Institute in the United States.

[4] Yan, J., Lin, M. 1988. Acoustic characteristics of the stress in Beijing trisyllables, *Fangyan* 1988.3, 227-237.

[5] Wang, J., Wang, L. 1993. The types of relative lengths of syllables in polysysllabic words in Putonghua. *Zhongguo Yuwen* 223, 112-116.