Porta Linguarum, Issue 42, June 2024 127-145

Cross-linguistic perception of Spanish intonation by Chinese speakers: Effects of linguistic experience and prosodic features

Peizhu Shang
Beijing Institute of Technology

Paolo Roseano
Universidad Nacional de Educación a Distancia & University of South Africa

Wendy Elvira-García
Universitat de Barcelona

Received: 2023-01-04 / Accepted: 2024-03-12
DOI: *https://doi.org/10.30827/portalin.vi42.27042*
ISSN paper edition: 1697-7467, ISSN digital edition: 2695-8244

ABSTRACT: Prior research suggests that Chinese speakers, due to their extensive experience with a tonal language, have an enhanced ability to discern pitch contrasts. This study investigates whether this aptitude extends to perceiving intonation in second language (L2) Spanish, and how it interacts with intonation and stress patterns. 43 native Spanish speakers and 75 Chinese learners of Spanish participated in a forced-choice perceptual test. Participants were asked to identify the sentence type of stimuli, which were created using the gating method. The results revealed that native speakers consistently outperformed Chinese learners in mapping diverse pitch contours into intonation categories, notably those absent in the L2 prosodic system. However, Chinese learners exhibited significant improvements in intonation perception after hearing the initial F0 peak and the F0 downslope of the nuclear syllable, suggesting a positive transfer of pitch discrimination skills developed in their native speech. Additionally, the position of the final stressed syllable was found to influence intonation identification, with its effect varying depending on utterance length. These findings highlight the complex interplay of cross-linguistic intonation perception and emphasize the importance of linguistic background and prosodic context in assessing the perceptual performance of L2 listeners.
Keywords: intonation contours, cross-linguistic perception, Chinese learners of Spanish, linguistic experience, stress pattern

Percepción interlingüística de la entonación del español por sinohablantes: Efectos de la experiencia lingüística y los rasgos prosódicos

RESUMEN: Investigaciones previas sugieren que los sinohablantes, debido a su amplia experiencia con una lengua tonal, tengan una mejor capacidad para discernir contrastes tonales. Este estudio investiga si esta aptitud se extiende a la percepción entonativa en el español como segunda lengua (L2) y cómo interactúa con patrones entonativos y acentuales. 43 hablantes nativos de español y 75 aprendices chinos de español participaron en un test perceptivo de elección forzada. Se pidió a los participantes que identificaran el tipo oracional de los

estímulos, los cuales fueron creados usando el método de *gating*. Los resultados revelaron que los hablantes nativos superaron consistentemente a los aprendices chinos en la categorización de contornos entonativos, especialmente aquellos ausentes en el sistema prosódico de la L2. Sin embargo, los aprendices chinos mostraron mejoras significativas en la percepción entonativa después de escuchar el pico inicial y la bajada de F0 de la sílaba nuclear, sugiriendo una transferencia positiva de habilidades discriminativas tonales desarrolladas en su habla nativa. Además, se encontró que la posición de la última sílaba acentuada influye en la identificación entonativa, variando su efecto dependiendo de la longitud oracional. Estos hallazgos resaltan la compleja interacción de la percepción de la entonación interlingüística y enfatizan la importancia de los antecedentes lingüísticos y el contexto prosódico en la evaluación del rendimiento perceptivo de los oyentes de L2..
Palabras claves: contornos entonativos, percepción interlingüística, aprendices chinos de español, experiencia lingüística, patrón acentual

1. Introduction

In Peninsular Spanish, statements and yes/no questions can be ambiguous from a morphosyntactic point of view, therefore intonation is often used linguistically to help listeners distinguish sentence types (Face, 2007; Shang et al., 2021). Similar to other intonational languages such as English, Spanish intonation is realized by rising-falling pitch movements over utterances. Thus, changing the pitch pattern of an utterance can change the statement into question (Kapatsinski et al., 2017). In contrast, in tonal languages like Mandarin Chinese, the higher-level sentence intonation is expressed mainly by expanding or compressing the pitch range, because the surface fundamental frequency (F0) contour is primarily used to encode word-level lexical tone (Shang & Elvira-García, 2022; Wang et al., 2013; Yuan, 2011). These cross-linguistic differences in the prosodic system would influence the way first language (L1) speakers of a tonal language produce and perceive the intonation of a non-native intonational language.

In recent decades, within the Autosegmental-Metrical (AM) framework, there has been a significant advancement in instrumental descriptions of the intonation of second languages (L2). This development encompasses not only the production but also the perception of phonological elements and phonetic properties. However, research focusing on the perception of L2 intonation in the context of language contact between tonal and non-tonal languages has predominantly concentrated on Chinese learners of English (e.g., Grabe et al., 2003; Liu & Rodriguez, 2012; Ortega-Llebaria et al., 2017; Zhang et al., 2010). Conversely, empirical studies addressing the perception of Spanish intonation by Chinese learners, both from Taiwan and mainland China, remain limited, with notable contributions from Cortés Moreno (1997, 2001), Li (2020), Liu (2005), and Shang et al. (2022, 2024). The initial studies by Cortés Moreno focused on the challenges faced by L2 Taiwanese learners in identifying different sentence types in Spanish, such as declarative, interrogative, and exclamative forms. Subsequent research by Li and Shang et al. focused, respectively, on the perception of focal contrasts and the auditory processing of acoustic cues in Spanish statements and questions. Despite the variations in experimental methodologies and research aims, a consistent finding across these studies showed that Chinese L2 learners deviated from native Spanish speakers in perceiving target intonational events because their prior history of linguistic experience was different.

In addition to the influence of linguistic background, other prosodic factors such as the stress pattern of the utterance-final word may affect listeners' perceptual evaluation of target intonation categories (Garrido Almiñana, 2008; Shang et al., 2022, 2024). However, it is not yet clear whether the effect of stress is evident in a cross-linguistic listening environment. Therefore, the present paper aims to investigate the non-native perception of Spanish intonation patterns by L1 and Chinese L2 speakers of Spanish. Specifically, we are interested in examining how Chinese L2 learners from a native tonal language rely on different intonational cues to perceive question-statement contrasts in Spanish, and how their perceptual accuracy is influenced by their linguistic experience and the position of the last stressed syllable compared to Spanish L1 listeners.

1.1. Intonation perception in L1 and L2 Spanish

In Peninsular Spanish, statements and yes/no questions can differ in four intonational aspects: initial F0 peak, pitch movement in the medial position, pitch contour of the final stressed syllable, and the final boundary tone (Face, 2004). Specifically, statements typically present an F0 contour that rises from the onset of the stressed syllable and has its peak in the poststressed syllable. And after the last prenuclear syllable, the contour undergoes a progressive down-stepping till the end (Estebas-Vilaplana & Prieto, 2010; Hualde & Prieto, 2015). Yes/no questions are also known to have a rising pitch accent in the prenuclear position, but their initial F0 peak is higher and the rising movement starts at the offset of the accented syllable (Estebas-Vilaplana & Prieto, 2010; Face, 2007, 2008). After that, the F0 pattern begins to fall gradually until the final nuclear syllable and then shows a sharp rise at the end of the sentence. In addition to prior contrasts, Face (2004, 2005, 2007) noted other two subtler intonational variances between these sentence types: (a) an F0 rise in the medial position of statements, absent in yes/no questions – noticeable only in utterances with at least three stressed words, and (b) a slight rise in the final stressed syllable of statements as opposed to an F0 dip in yes/no questions.

Based on the differences described in Spanish statements and questions, Face (2005) employed a gating paradigm to investigate how different intonation patterns influence sentence type perception. His experimental results showed that Spanish L1 listeners were able to correctly perceive 95% of the sentence type after hearing the initial F0 peak, and this accuracy rate continued to increase after presenting the gate that contained the medial F0 pattern. Later, to compare the weight of each intonational cue, Face (2007) manipulated the height of several pitch points so that the intonational features presented in the utterance were in conflict. Results of this experiment showed that the final pitch movement was the strongest cue for distinguishing statements and questions, despite the heavy influence that the first peak and the median F0 pattern had on listeners' perceptual decisions.

Contrary to the sentence type perception in L1 Spanish, the non-native perception of Spanish intonation has been less extensively studied. Trimble (2013) examined the cross-linguistic perception of Spanish intonation using a gating paradigm and found a heavy reliance on the final pitch movement for sentence type identification. Specifically, 58% of English L2 speakers failed to distinguish Spanish statements from yes/no questions without the final rise/fall. However, a similar study with Chinese learners of Spanish (Li, 2020) showed higher perceptual accuracy (68.33%) based solely on the first F0 peak. The superior performance of Chinese speakers seems to support the notion that long-term experience with a tonal language may aid in listeners assimilating non-native tonal contrasts (Meng et al., 2020; Ortega-Llebaria et al., 2017). But despite this tonal advantage, Chinese L2 learners still lagged behind native Spanish speakers in utilizing the prenuclear peak for sentence differentiation (Li, 2020).

Furthermore, studies by Li (2020), Trimble (2013), and Zarate-Sandez et al. (2015) indicated that both Chinese and English learners of Spanish could achieve comparable accuracy to native speakers upon hearing the final pitch movement. This aligns with Gussenhoven and Chen's (2000) findings, where speakers of diverse languages (Chinese, Dutch, Hungarian) effectively used higher final pitch and later F0 peaks as cues for question perception in unfamiliar languages. A generally accepted explanation for this phenomenon is that such intonational forms, particularly the final rise/fall, can be universal in the relationship with the paralinguistic meanings and derived from some biologically determined codes (Gussenhoven, 2004). For instance, the Frequency Code posits that an increased F0 value expresses the speaker's uncertainty and hence is an interrogative marker, whilst a lower F0 signals assertiveness and therefore is associated with declarative encoding (Gussenhoven, 2004; Ohala, 1984).

1.2. The role of linguistic experience and prosodic features in intonation perception

Despite differences on deeper issues, speech perception theories show a strong agreement on the critical role of linguistic experience on cross-linguistic perception. The Perceptual Assimilation Model (PAM) and its extension to L2 learning (PAM-L2) posit that the perception of L2 segments can be implicitly or explicitly constrained by L1 phonological categories and phonetic properties (Best & Strange, 1992; Best & Tyler, 2007). Likewise, the PAM framework developed for suprasegmentals (PAM-S) suggests that the auditory performance of L2 listeners reflects the specific properties of their L1 phonological and phonetic systems (Alexander & Wang, 2016; So & Best, 2010). Such theoretical claims have been supported by studies from the prosodic and neurophysiological points of view. For instance, it has been reported that Chinese listeners' perception of the native or non-native tone (Chang et al., 2017), intonation (Ortega-Llebaria et al., 2017) and stress (Meng et al., 2020) is enhanced by their native tonal language experience. Particularly, Ortega-Llebaria et al. (2017) examined the perception of English words using the prime-target pairs and found that Chinese learners detected faster the F0 mismatches than English listeners, suggesting that the long-term experience with a tonal language can shape the way how listeners perceive the tonal contrasts inherent to a given language. The effect of tonal language enhancement on pitch perception has been found in a vast majority of studies, including the perception of local tonal events of second or third languages. For instance, Wiener and Goss (2018) investigated the perception of Japanese pitch accents by English speakers using an ABX discrimination task and found that the English group apparently improved their accuracy after taking a 15-week L2 Mandarin tone course.

On the other hand, acoustic analyses of Spanish intonation showed that the stress pattern in the utterance-final position may also modulate the evolution of the final pitch movement depending on the type of melodic patterns (Fant, 1984; Garrido Almiñana, 2008;

Navarro-Tomás, 1974; Shang, 2022a). The rising-falling intonation contours were reported to be the most dependent on the location of the final stressed syllable of an utterance, while the shape of falling patterns did not seem to be dependent at all (Garrido Almiñana, 2008). Despite the potential influence of stress on the production of intonation contours, it remains unclear whether the position of the last stressed syllable affects the recognition of statements and questions by L1 and L2 listeners. Recent cue-weighting studies on Spanish intonation observed that words with stress on the final syllable were difficult to perceive by Mandarin L2 learners in terms of intonation perception, due to the conflict of functions in f0 encodings in utterance-final positions (Shang et al., 2022, 2024). However, these studies focused on listeners' sensitivity to changes in acoustic cues (F0, duration, and intensity), whether and how the stress pattern in the utterance-final position might affect cross-linguistic perception of categorical elements of Spanish intonation remain to be explored.

1.3. The present study

Considering the discussions outlined above, it is evident that the cross-linguistic perception of intonation is subject to a variety of influencing factors. Therefore, the primary aim of this study is twofold: (a) to explore the impact of these factors on the perception of Spanish intonation, and (b) to delineate the perceptual challenges faced by Chinese L2 learners in contrast with native Spanish speakers. To achieve these objectives, we conducted a gating experiment focusing on the categorical perception of intonational patterns and contrastive structures in Spanish statements and yes/no questions. Our examination specifically revolves around three dimensions: (1) the shape of intonation contours; (2) the influence of being a native speaker of a tone language; and (3) the position of the final stressed syllable. This research is driven by three core questions regarding the cross-linguistic perception of Spanish intonation, particularly pertaining to the Chinese-Spanish language pairing, which have not been thoroughly addressed in prior studies:

- RQ1: How do Chinese L2 learners compare to native Spanish speakers in their perception of varying intonational contours in statements and yes/no questions?
- RQ2: In what ways does linguistic background shape the categorical perception of Spanish intonation? Specifically, does extensive experience with a tonal language enhance L2 learners' ability to perceive non-native sentence types?
- RQ3: Does the stress pattern in the final position of an utterance influence the cross-linguistic perception of Spanish sentence types?

2. Methodology

2.1. Materials and recordings

We used four pairs of broad focus statements and yes/no questions to generate the stimuli (see Table 1). The statement and the yes/no question of each pair were morphosyntactically identical. The four sentence matches differed in their length and therefore in their number of stressed words. For each length, we included two stress patterns for the utterance-final word: penultimate stressed syllable (paroxytone) and final stressed syllable (oxytone). The DCT (Discourse Completion Task) technique was used to elicit a natural production of the statements and the yes/no questions (Félix-Brasdefer, 2010). A female native speaker of Spain (age at recording: 31) was asked to give her response at a normal speaking rate after presenting each situational context. The recordings took place in a quiet room using the microphone Rode Smartlav+ connected to the interface of Scarlett. Audio files were digitized at a sampling rate of 44.1 kHz and with a quantization precision of 16 bits.

Table 1. *Categorization of the recorded utterances*

Sentence type	N. of stress	Final stressed word	Stimulus
Statement	1	Oxytone	Alcalá.
Statement	1	Paroxytone	Sevilla.
Statement	2	Oxytone	Viene a Alcalá.
Statement	2	Paroxytone	Viene a Sevilla.
Yes/no question	1	Oxytone	¿Alcalá?
Yes/no question	1	Paroxytone	¿Sevilla?
Yes/no question	2	Oxytone	¿Viene a Alcalá?
Yes/no question	2	Paroxytone	¿Viene a Sevilla?

2.2. Stimulus creation

The recordings described above were used to create the stimuli. In order to do so, a gating method was employed based on the rules proposed by Face (2005, 2007). Each of the four sentences with two stressed words was split into four gates. Contrary to what is usually done, we did not use segmental information to create the steps, instead, we used intonational events which contain the three intonational differences between the statement and the yes/no question. As a result, gates coincided with known intonational events in Spanish phonology and therefore recordings with one and two stressed words had a different number of steps.

For sentences with two stressed words, Gate 1 consisted of the F0 contour during the first stressed syllable “Vie” where the statement showed a slight rise starting near the onset of the stressed syllable while the yes/no question no (see Figure 1). Gate 2 included the initial F0 peak, and it may sometimes contain an adjacent syllable of the following word because of the synaloepha. This peak of the second gate was reported to be the first strong differential cue in identifying the two-word sentences, given that questions have a different prenuclear pattern than statements in Spanish (Face, 2007, 2011). That is also the case in our study, for example, in Figure 1, the yes/no question has a 71 Hz higher F0 peak (335 Hz) than the statement (264 Hz). Gate 3 begins after the first peak and before the utterance-final syllable. Hence, for sentences ending with a paroxytone word, gate 3 included the F0 contour of the final stressed syllable, which unequivocally differentiates the yes/no question (rising contour) from the statement (low-falling contour). In the case of utterances ending with an oxytone word, gate 3 did not contain a stressed syllable. This means that gate 3 has more cues for the correct recognition in final paroxytone sentences than in oxytone ones. We decided to keep this division in order to create the same number of gates for the two stress patterns. This difference in gate 3 also allows us to explore the role of the final stressed syllable in perceiving question-statement contrasts. The last gate (gate 4) of the two-word sentences consisted of the intonation patterns of the final syllable which traditionally have been referred to as the last but most typical cue for signalling the statement and the yes/no question.

Regarding the sentences with one stressed word, we created two gates. The boundary of the gates was established just before the utterance-final syllable (see Figure 2). Similar to the longer sentences, the gating division of the two stress patterns in the one-word sentences was slightly different. The paroxytone word differed from the oxytone word in including the F0 contour of the stressed syllable in gate 1 (which was a relatively high-rising contour in statements whereas a low-falling contour in yes/no questions). The last gate (gate 2) was comprised of the F0 contours of the whole word, including the typical final pitch movement of statements and questions.

In both cases, the division of each gate for the eight sentences was realized using Praat (Boersma & Weenink, 2020) without modifying any acoustic signal. The final stimuli were generated by successively accumulating the first gate, the first and the second gate, and so on until involving the entire utterance. In sum, a total of 24 stimuli were created for the gating experiment. Each stimulus was separated and saved as a new sound file, with 500 milliseconds of silence before and after the speech.

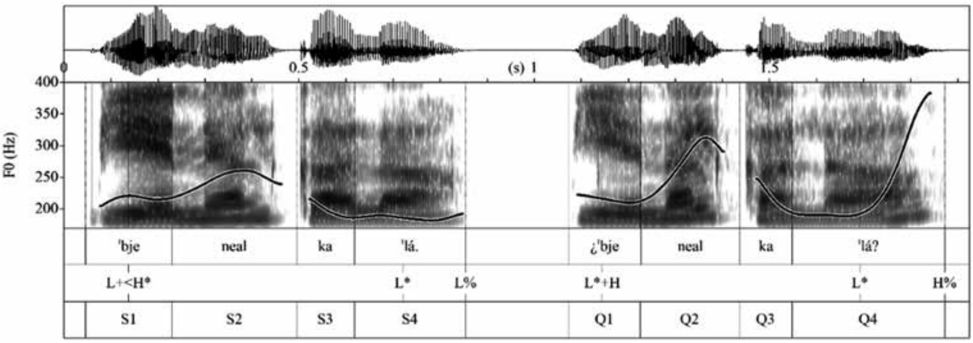

Figure 1. *Intonation contours of the sentence pair with two stressed words aligned with the syllables and the gates (S1, gate 1 of the statement, Q1, gate 1 of the yes/no question, and so on for the rest of the abbreviations)*

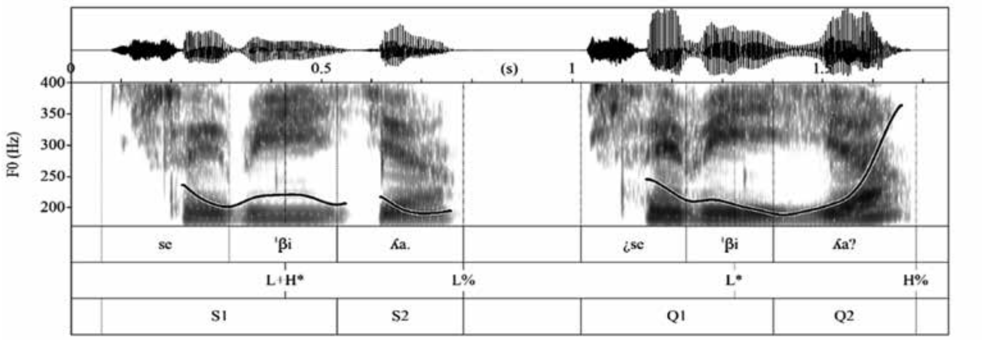

Figure 2. *Intonation contours of the one-word statement and yes/no question aligned with the syllables and the gates*

2.3. Procedure

The data was gathered using the online survey software Alchemer. The survey consisted of two sections. The first aimed to collect information on the socio-demographic and linguistic background of the participants. The second contained the stimuli of the gating experiment. The survey was distributed separately in Chinese and Spanish via e-mail and social media platforms. Participants were required to listen to the stimuli using the earphones in a quiet place and they were instructed to listen to every audio once. Prior to the formal task, participants performed a practice trial to get familiar with the procedure. Additionally, given the difficulty that a gating test can be, the text of the entire utterance was displayed without punctuation marks. The task of the gating test was identifying whether the stimulus they heard was “Statement” or “Yes/no question” by clicking a button on the screen. To avoid listeners’ response bias, the perceptual choices and speech stimuli were set in random order.

2.4. Participants

Data cleaning was performed before analysis. Data from two Spanish listeners were excluded because their age was over 60. Similarly, Chinese speakers who self-declared to have learned Spanish before the age of 16 were removed from this study to minimize the effect of age of acquisition. Thus, there were forty-three Spanish L1 listeners and seventy-five Chinese L2 learners of Spanish in the present study. Of the Chinese subjects, 22 had a B1 level in Spanish, 27 had a B2 level in Spanish, and 26 had a C1 level in Spanish. The ages of all participants ranged from 18 to 59 years with a mean age of 28.77 (*SD* = 8.35). Among them, 28 were male and 90 were female. None of the individuals reported any history of hearing or communication problems at the time of testing.

2.5. Statistical analysis

To examine the research questions proposed in this study, a binary logistic regression was carried out using the *glm* function in the base R package (R Core Team, 2020). The binary outcome variable "perceptual accuracy" was created based on listeners' responses to each stimulus. The true and false answers were coded as 1 and 0, respectively. Listener group (Chinese vs. Spanish), sentence type of the stimuli (statement vs. yes/no question), and stress pattern (paroxytone vs. oxytone) were treated as binary variables, while gate ($1 < 2 < 3 < 4$ for two-word sentences; $1 < 2$ for one-word sentences) were read as discrete ordinal variables. For the continuous predictors of stimulus order, a z-transformation was performed based on the mean and standard deviation before entering the model.

To find the best-performing model, a stepwise backward elimination was carried out using the *stepAIC* function from the *MASS* package to iteratively exclude the least contributive predictors (Venables & Ripley, 2002). When the factors failed to achieve the .05 level of significance, they were removed from the model until all effects were statistically significant for the prediction. For further interaction analysis, *emmeans* objects were created using the *contrast* function to estimate the contrast of interests (Lenth et al., 2018).

3. Results

3.1. Results of the sentences with two stressed words

Since the effects of stress type and stimulus order were statistically irrelevant, they were not included in the final model. The results of the likelihood ratio test revealed a highly significant effect between the perceptual accuracy and the variables of listener group, sentence type and gate [$\chi^2(13) = 527.66$, $p < .0001$]. In general, Table 3 indicates that listeners were more accurate in perceiving statements than yes/no questions. However, the effect of sentence type was not distributed evenly across the four gates. The significant interaction between sentence type and gate [$\chi^2(3) = 35.09$, $p < .0001$], as shown in Table 2, demonstrated that the accuracy of perceiving statements was significantly higher than yes/no questions only in gate 1. More precisely, by transforming the regression coefficient in Table 2 into odds ratios (OR), we found that statements were 3.69 times more likely to be correctly perceived than yes/no questions in the first gate [OR = exp(1.3053) = 3.69, $p < .0001$].

Table 2. *Results of the regression model fitted for sentences with two stressed words*

Fixed effect	*Coef*	*SE*	*Z*	*P*
(Intercept)	0.7770	0.209	3.71	<.001**
Sentence type*Gate (Q1 vs. S1)	-1.3053	0.195	-6.70	<.0001***
Sentence type*Gate (Q2 vs. S2)	0.3676	0.222	1.65	>.05
Sentence type*Gate (Q3 vs. S3)	-0.0825	0.287	-0.29	>.1
Sentence type*Gate (Q4 vs. S4)	-1.4066	1.124	-1.25	>.1
Group*Gate (SP1vs. CH1)	0.3142	0.202	1.55	>.1
Group*Gate (SP2 vs. CH2)	1.5251	0.284	5.37	<.0001***
Group*Gate (SP3 vs. CH3)	1.9698	0.478	4.12	<.0001***
Group*Gate (SP4 vs. CH4)	13.4462	294.463	0.05	>.1
Gate*Group (CH1 vs. CH2)	0.8321	0.173	4.80	<.0001***
Gate*Group (CH2 vs. CH3)	0.8412	0.195	4.30	<.0001***
Gate*Group (CH3 vs. CH4)	2.7613	0.582	4.74	<.0001***
Gate*Group (SP1 vs. SP2)	2.0430	0.303	6.74	<.0001***
Gate*Group (SP2 vs. SP3)	1.2859	0.521	2.47	>.05
Gate*Group (SP3 vs. SP4)	14.2381	294.463	0.05	>.1

Note: S1, gate 1 of the statement; Q1, gate 1 of the yes/no question; SP1, Spanish native listeners in gate 1; CH1, Chinese learners of Spanish in gate 1; and so on for the rest of the abbreviations

Table 3. *Descriptive statistics of the perceptual accuracy (%) after each of the gates of the sentences with two stressed words*

Gate	Statement		Yes/no question		Combined total	
	CH	SP	CH	SP	CH	SP
1	62.67%	69.77%	31.33%	38.37%	47.00%	54.07%
2	63.33%	87.21%	70.00%	93.02%	66.67%	90.12%
3	84.00%	95.35%	80.67%	98.84%	82.33%	97.09%
4	99.33%	100.00%	97.33%	100.00%	98.33%	100.00%

Note: CH, Chinese learners of Spanish; SP, Spanish native listeners

On the other hand, the significant interaction between listener group and gate [$\chi^2(3) = 18.20$, $p < .0001$] showed that Spanish L1 and L2 listeners differed in the way they perceived each gate, although both listener groups could improve the accuracy with accumulating the intonational cues. Specifically, in the first gate, Chinese and Spanish listeners showed no significant difference in perceptual accuracy, as revealed by Table 2. However, in the second gate, Spanish listeners were 4.60 times more likely than Chinese learners to correctly identify sentence types [OR = exp(1.5251) = 4.60, $p < .0001$], although the Chinese group also showed a significant improvement from the first to the second gate [OR = exp(0.8321)

= 2.30, $p < .0001$]. The apparent increase of accuracy shown by both listener groups in gate 2 was cued by the presence of the initial F0 peak, which was higher in the yes/no question and lower in the statement.

Further, from gate 2 to gate 3, the F0 downslope after the highest peak and before the final pitch change was referred to as the second intonational cue for question-statement identification (with the yes/no question having a stepper F0 slope than the statement). Compared to the second gate, the likelihood of accurately perceiving the sentence type in gate 3 was 2.32 times greater for Chinese learners [OR = exp(0.8412) = 2.32, $p < .0001$]. Although Spanish listeners did not show significant progress after hearing gate 3, they were still more accurate than L2 learners in the identification task (see Table 3). After gate 3, when the entire utterance was released, both listener groups were able to perceive the sentence type with a nearly 100% accuracy rate, suggesting that the final F0 rise may be a universal auditory cue for the activation of interrogative meanings.

Finally, as opposed to gate 3, the odds of accurately perceiving the sentence type in gate 4 were significantly (15.82 times) greater for Chinese learners [OR = exp(2.7613) = 15.82, $p < .0001$. While for Spanish L1 listeners, it may seem that the final F0 contour was not important as the initial peak for the perception of sentence types, because their increase in accuracy was not significant after being presented with gate 4 (see Table 2). However, this impression has proved to be false by Face (2007, 2011), which showed that the final pitch movement was the strongest cue for native Spanish listeners in determining the sentence type that can override any other previous cues of the sentence. To further confirm the strength of the final F0 contour in cross-linguistic perception, sentences with one stressed word were analyzed in Section 3.2 because of their advantages in reducing the interference of other intonational cues in perception.

3.2. Results of the sentences with one stressed word

Data from the one-word sentences was analyzed by building a new GLM model with a similar fitting process. The significance test of the model showed a highly significant effect between the perceptual accuracy of the one-word sentences and the variables of sentence type, gate, listener group and stress pattern [$\chi^2(5) = 437.99$, $p < .0001$]. Considering the effect of sentence type, Table 4 indicates that listeners had significantly lower accuracy in perceiving yes/no questions rather than statements. All else being equal, the yes/no question had an 85% lower likelihood of being correctly identified compared to the statement over the utterance [OR = exp(-1.8959) = 0.15, $p < .0001$]. The result is in parallel with our previous findings of the longer sentences, suggesting that there might be a hierarchy of difficulties in perceiving different sentence types. Moreover, despite there being no significant interaction of sentence type with gate in one-word sentences, the perceptual differences between the statement and the question were mainly shown on the first gate (see Table 5). From gate 1 to gate 2, all listeners experienced a highly significant increase in the likelihood that a given utterance was correctly perceived and achieved a nearly 100% accuracy rate. This result provides evidence for the important role of the final F0 contour in both L1 and L2 perception, overthrowing the view that the final pitch movement had little or less strength compared with other cues in identifying the sentence type.

Furthermore, the positive coefficient of listener group (see Table 4) indicates that Spanish native listeners were significantly more accurate than Chinese L2 learners in the overall identification of one-word sentences. Apart from this distinction, a unique result was obtained for the variable of stress pattern. Table 4 shows that the paroxytone word had a 2.32 times higher likelihood of being correctly identified compared to the oxytone word [OR = exp(0.8422) = 2.32, $p < .01$]. This finding seems to suggest that the location of the final stressed syllable may also play a role in perceiving the one-word sentences of Spanish. However, as revealed by Table 5, the perceptual advantage of the paroxytone word was mainly restricted to the first gate. The interpretation of this result may be related to the different phonological components involved in the first gate of the two stressed words, which we will discuss in Section 4.

Table 4. *Results of the regression model fitted for sentences with one stressed word*

Fixed effect	*Coef*	*SE*	*Z*	*P*
(Intercept)	0.7770	0.209	3.71	<.001**
Sentence type (Statement vs. Question)	-1.8959	0.221	-8.57	<.0001***
Gate (2 vs.1)	5.2728	0.597	8.83	<.0001***
Group (SP vs. CH)	0.6919	0.220	3.15	<.01*
Stress (Paroxytone vs. Oxytone)	0.8422	0.217	3.89	<.01*

Note: CH, Chinese learners of Spanish; SP, Spanish native listeners

Table 5. *Descriptive statistics of the perceptual accuracy (%) after each of the two gates of the sentences with one stressed word*

Gate	Stress pattern	Statement		Yes/no question		Combined total	
		CH	SP	CH	SP	CH	SP
1	Paroxytone	78.67%	95.35%	34.67%	51.16%	57.33%	73.26%
1	Oxytone	65.33%	72.09%	25.33%	32.56%	45.33%	52.33%
2	Paroxytone	100.00%	100.00%	98.67%	100.00%	99.33%	100.00%
2	Oxytone	98.67%	100.00%	98.67%	100.00%	99.33%	100.00%

Note: CH, Chinese learners of Spanish; SP, Spanish native listeners

4. Discussion

This study examines the cross-linguistic perception of Spanish intonation by Chinese tonal speakers and evaluates the factors that may guide listeners' prosodic identification of target sentence types. We start to consider the first research question and argue primarily that the cross-linguistic perception of Spanish intonation was strongly influenced by the different intonational cues to which L1 and L2 listeners were differently sensitive. Evidence for this is that both Spanish and Chinese listeners could progressively improve their perceptual accuracy by accumulating the number of gates in the utterance. This finding is consistent with the studies of Face (2005, 2007) that all intonational cues contributed at least to some extent to the perception of question-statements contrasts. However, as revealed by the significant interaction between listener group and gate, L1 and L2 listeners differed in the ability to use distinct intonational cues to help increase the likelihood that a given utterance is correctly identified. This difference is particularly significant in perceiving the initial F0 peak and the F0 downslope of the nuclear syllable of the longer sentences, while in the first and the final gate, the two listener groups showed statistically comparable performances. Similar challenges in perceiving L2 Spanish intonation, as documented by Li (2020), Shang (2022b), and Trimble (2013), likely stem from the different intonational forms inherent in the native prosodic systems of learners. The F0 contours in a tone language such as Mandarin are used primarily to encode lexical meanings rather than for discriminating intonation types (Yuan, 2011). Therefore, Chinese speakers do not necessarily rely on the variation of pitch contours to make clear the sentence type (Shang & Elvira-García, 2022). But despite the difference in intonation mechanism, Chinese learners of this study achieved roughly 67% accuracy rate as soon as they were presented with the first F0 peak.

This finding is somewhat surprising in comparison with the results of Trimble (2013) that more than half of the English L2 learners failed to distinguish Spanish statements and questions without the final pitch movement. However, in response to the second research question, the superior performance of Chinese learners was expected, which is consistent with previous literature that listeners with tonal language experience were more sensitive than non-tonal language listeners in perceiving the contrastive tone patterns (Meng et al., 2020; Ortega-Llebaria et al., 2017; Ortega-Llebaria & Wu, 2021). The enhanced perception exhibited by Chinese learners in this study can be interpreted as a positive transfer of their pitch processing abilities developed in L1 speech. Speakers of tonal languages share the neural architecture that processes signals with musicians and therefore, were associated with higher auditory perceptual performance and enhanced general cognitive abilities in perceiving tonal features or complex music than speakers of a non-tonal language and nonmusicians (Bidelman et al., 2013).

Although the initial F0 peak and the F0 downslope are of great importance in perceiving longer sentences, it is only when the final pitch movement was presented that Chinese learners could achieve an accuracy level similar to that of the L1 listeners. The significant progress of Chinese learners after hearing the final gate corroborates previous findings (Li, 2020; Shang, 2022; Trimble 2013), showing that the final rise/fall was a strong intonational cue for non-native listeners in decoding the sentence type. As for Spanish native listeners, it could seem that the final intonational pattern played a minor role in identifying longer sentences. Yet, by using one-word sentences as stimuli, we refute this false impression by showing that the seemingly lower weighting of the final gate was because there were apparent prosodic makings (i.e., initial F0 peak) before the utterance-final end which could lead Spanish L1 listeners to make a confident perceptual decision without the presence of the final intonational cue. In other words, the final pitch movement still was the greatest contributing factor to the likelihood that a given utterance is accurately recognized, either for L1 or L2 listeners. This finding is in line with the results of Face (2007) that the final F0 pattern was the strongest cue for signalling the sentence type that had the strength to change listeners' auditory decisions made with earlier contradictory cues. The consistent weighting of the final pitch movement across language groups also seems to support the claim that there may be a language-universal mechanism for the relationship between intonational forms and pragmatic meanings in speech (Gussenhoven & Chen, 2000; Gussenhoven, 2004).

As regards the third research question, our results partially confirm that the auditory performance of Chinese learners was influenced by the prosodic structure of the speech materials, i.e., the position of the final stressed syllable, in addition to the linguistic experience. For example, it is found that sentences ending with a paroxytone word were more likely to be correctly perceived than the ones ending with an oxytone word before presenting the final syllable. As explained earlier, the perceptual difference probably is attributed to that the first gate of the paroxytone word was divided including the stressed syllable, while that of the oxytone word did not include it. In other words, the paroxytone word had a distinctive cue in the first gate to help increase its perceptual accuracy, since the statement in Spanish is described to have an F0 rise during the final stressed syllable, as opposed to the yes/no question that is realized with a low F0 throughout this syllable (Face, 2007). However, note that the effect of stress pattern was apparent only for the one-word sentences (given that the nuclear pitch accent is L+H* for statements and L* for questions), while in utterances with two stressed words, no significant difference in accuracy was observed depending on the stress pattern (because the nuclear pitch accent is L* for both questions and statements). A similar finding was reported by Face (2005) which indicated that the F0 contour of the final stressed syllable was a weaker cue for discriminating the utterances with three stressed words, but it was of great importance for rating the naturalness of yes/no questions.

Lastly, in line with earlier works on intonation perception (Face, 2005, 2007; Trimble, 2013; Li, 2020), our study demonstrated that L1 and L2 listeners of Spanish generally were biased towards identifying an utterance as a statement, particularly when there is no apparent interrogative cue involved in the gate. This parallel inclination between different language groups could be interpreted as a preference for unmarked forms in communicative interaction (Grosser, 2011). Given that statements are the most neutral and unmarked sentence modality that carries the least communicative burden in most human languages, it would be, of course, more easily to be recognized than questions in human speech.

5. Conclusions and implications

The present study employed a gating paradigm to explore the perception of Spanish intonation patterns among native Spanish speakers and Chinese L2 learners, focusing on how their auditory performance is influenced by linguistic experience and prosodic structures of the stimuli. The results, on the one hand, underscore the significance of linguistic experience in processing target intonational events, particularly in discerning question-statement contrasts. Chinese L2 learners, compared to native Spanish speakers, showed variance primarily in perceiving intonational contours absent in their native prosodic system, such as the initial F0 peak and the F0 downslope of the nuclear syllable. However, there was a notable consistency in the auditory weighting of the final pitch movement across both L1 and L2 Spanish listeners. This perceptual similarity across linguistic backgrounds supports theories suggesting the existence of universal biological codes in linguistic meaning (Chen, 2009; Chen, Gussenhoven, & Rietveld, 2004), indicating a blend of universal and language-specific influences on intonation processing. Furthermore, relative to speakers of non-tonal languages like English (Trimble, 2013), Chinese L2 learners displayed a more native-like perception of the prenuclear stretch. This enhanced perceptual ability is attributed to their extensive tonal language experience, which likely bolsters their cognitive and perceptual skills in processing contrastive pitch patterns, especially those with specific curvatures or directional changes. Additionally, the study reveals that cross-linguistic perception is influenced not only by linguistic background but also by factors like the location of the final stressed syllable and the utterance length.

While this study provides valuable insights into the perception of Spanish intonation by Chinese L2 learners, it is important to acknowledge certain limitations. Firstly, the auditory materials used encompassed a limited range of sentence types, restricting a more comprehensive evaluation of intonation perception abilities. This limitation may potentially influence the generalizability of our findings to other sentence structures in Spanish. Secondly, the study did not account for the individual properties of Chinese L2 learners, such as varying levels of L2 proficiency, pragmatic competence, or musical ability. These factors might significantly impact learners' perception and processing of intonational cues, suggesting that future research should consider these variables for a more nuanced understanding. Lastly, the auditory materials were developed from the speech output of a single Spanish female speaker, raising the question of whether speaker gender or voice quality might differentially influence the intonation perception among L2 listeners of different genders. This aspect could be a critical factor in understanding the cross-linguistic perception of intonation and warrants further exploration to determine its impact on auditory processing in a second language context.

5.1. Pedagogical implications and future directions

The findings of this study have critical implications for language education, particularly in the context of teaching Spanish intonation to learners from mainland China. Understanding the specific areas where Chinese L2 learners of Spanish excel or face perceptual challenges can guide the development of targeted teaching strategies and materials. For instance, recognizing the difficulties Chinese learners encounter with certain Spanish intonational contours that are absent in Mandarin can lead to the development of specialized training modules. These modules could focus on enhancing the perception and production of these specific intonation patterns.

Further, the insights into the perceptual similarities and differences between native Spanish speakers and Chinese learners of Spanish can guide the creation of more refined teaching methods. Such methods should capitalize on the strengths and address the unique challenges of these learners. In this context, innovative approaches like using embodied gestures and movements (e.g., Baills et al., 2019; Yuan et al., 2019), audiovisual speech analysis software (e.g., Bengrait, 2018), and musical training activities (e.g., Moradi & Shahrokhi, 2014) could be particularly effective. These techniques can help in making the learning process more interactive and engaging, thereby improving the learners' ability to both perceive and produce Spanish intonation accurately.

Looking ahead, a natural extension of this research would be to develop and test these pedagogical methods and techniques in real classroom settings. This would not only validate their effectiveness in enhancing Spanish intonation acquisition among Chinese learners but also contribute to the broader field of L2 speech acquisition. Ultimately, the goal is to draw greater attention to the often-neglected aspect of intonation in language teaching and learning, particularly in the Chinese-Spanish linguistic context. By doing so, educators and learners can gain a deeper appreciation of this linguistic phenomenon, leading to more comprehensive and effective language education.

FUNDING & ACKNOWLEDGMENTS

This study received funding from the Beijing Institute of Technology Research Fund Program for Young Scholars (No. 3240012222308) and the Ministry of Education University-Industry Collaborative Education Program of China (No. 230824085307241), awarded to the first author (P.S.). Peizhu Shang of the Beijing Institute of Technology is the corresponding author for this paper.

We extend our sincere gratitude to the participants of this study, the anonymous reviewers, and the editor for their invaluable contributions and insights.

6. References

Alexander, J., & Wang, D. Y. (2016). The perceptual assimilation model for suprasegmentals and cross-language lexical-tone identification. *The Journal of the Acoustical Society of America*, *139*(4), 2161–2161. https://doi.org/10.1121/1.4950403

Baills, F., Suárez-González, N., González-Fuente, S., & Prieto, P. (2019). Observing and producing pitch gestures facilitates the learning of Mandarin Chinese tones and words. *Studies in Second Language Acquisition*, *41*(1), 33–58. https://doi.org/10.1017/S0272263118000074

Bengrait, N. (2018). *The use of speech analysis software for the teaching of English intonation* [Unpublished doctoral dissertation]. University Badji Mokhtar-Annaba.

Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics*, *20*(3), 305–330. https://doi.org/10.1016/S0095-4470(19)30637-0

Best, C. T., & Tyler, M. (2007). Nonnative and second-language speech perception. In M. Murry J. & B. Ocke-Schwen (Eds.), *Language Experience in Second Language Speech Learning: In honor of James Emil Flege* (pp. 13–34). John Benjamins.

Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: evidence for bidirectionality between the domains of language and music. *PloS One*, *8*(4), 1–11. https://doi.org/10.1371/journal.pone.0060676

Boersma, P., & Weenink, D. (2020). *Praat: doing phonetics by computer* (Version 5.3.82) [Computer software]. http://www.praat.org/

Chang, Y. S., Yao, Y., & Huang, B. H. (2017). Effects of linguistic experience on the perception of high-variability non-native tones. *The Journal of the Acoustical Society of America*, *141*(2), EL120–EL126. https://doi.org/10.1121/1.4976037

Chen, A. (2009). Perception of paralinguistic intonational meaning in a second language. *Language Learning*, *59*(2), 367–409. https://doi.org/10.1111/j.1467-9922.2009.00510.x

Chen, A., Gussenhoven, C., & Rietveld, T. (2004). Language-specificity in the perception of paralinguistic intonational meaning. *Language and Speech*, *47*(4), 311–349. https://doi.org/10.1177/00238309040470040101

Cortés Moreno, M. (1997). Sobre la percepción y adquisición de la entonación española por parte de hablantes nativos de chino. *Estudios de Fonética Experimental*, *9*, 67–134.

Cortés Moreno, M. (2001). Percepción y adquisición de la entonación española en enunciados de habla espontánea: el caso de los estudiantes taiwaneses. *Estudios de Fonética Experimental*, *11*, 90–119.

Estebas-Vilaplana, E., & Prieto, P. (2010). Castilian Spanish intonation. In P. Prieto & P. Roseano (Eds.), *Transcription of intonation of the Spanish language* (pp. 17–48). Lincom Europa.

Fant, L. (1984). *Estructura informativa en español: estudio sintáctico y entonativo (Acta Universitatis Upsaliensis)*. Uppsala University.

Face, T. L. (2004). The intonation of absolute interrogatives in Castilian Spanish. *Southwest Journal of Linguistics*, *23*(2), 65–80.

Face, T. L. (2005). F0 peak height and the perception of sentence type in Castilian Spanish. *Revista Internacional de Lingüística Iberoamericana*, *3*(2), 49–65.

Face, T. L. (2007). The role of intonational cues in the perception of declaratives and absolute interrogatives in Castilian Spanish. *Estudios de Fonética Experimental*, *16*, 186–225. https://raco.cat/index.php/EFE/article/view/140052

Face, T. L. (2008). *The intonation of Castilian Spanish declaratives and absolute interrogatives*. Lincom Europa.

Face, T. L. (2011). *Perception of Castilian Spanish intonation: Implications for intonational phonology*. Lincom Europa.

Félix-Brasdefer, J. C. (2010). Data collection methods in speech act performance: DCTs, role plays, and verbal reports. In A. Martínez-Flor & E. Usó-Juan. (Eds.), *Speech Act Performance* (pp. 41-56). John Benjamins.

Garrido Almiñana, J. M. (2008). *Modelling Spanish intonation for text-to-speech applications* [Doctoral dissertation, Universitat Autònoma de Barcelona]. TDX. http://hdl.handle.net/10803/4885

Grabe, E., Rosner, B. S., García-Albea, J. E., & Zhou, X. (2003). Perception of English intonation by English, Spanish, and Chinese listeners. *Language and Speech*, *46*(4), 375-401. https://doi.org/10.1177/00238309030460040201

Grosser, W. (2011). On the acquisition of tonal and accentual features of English by Austrian learners. In A. Leather & J. & J (Eds.), *Second language speech* (pp. 211–228). De Gruyter Mouton.

Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge University Press.

Gussenhoven, C., & Chen, A. (2000). Universal and language-specific effects in the perception of question intonation. *Proceedings of the 6th International Conference on Spoken Language Processing* (pp. 91–94). Beijing, China.

Hualde, J. I., & Prieto, P. (2015). Intonational variation in Spanish: European and American varieties. In F. Sonia & P. Pilar (Eds.), *Intonation in Romance* (pp. 350–391). Hardcover.

Kapatsinski, V., Olejarczuk, P., & Redford, M. A. (2017). Perceptual learning of intonation contour categories in adults and 9- to 11-year-old children: Adults are more narrow-minded. *Cognitive Science*, *41*(2), 383–415. https://doi.org/10.1111/cogs.12345

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). Package “Emmeans”. R Package Version 4.0-3. http://cran.r-project.org/package=emmeans

Li, W. (2020). *Producción y percepción de la entonación del español por los sinohablantes: patrones entonativos prototípicos de la declarativa neutra, la interrogativa absoluta neutra y la declarativa con foco estrecho* [Doctoral dissertation, Universitat Autònoma de Barcelona]. TDX. http://hdl.handle.net/10803/670612

Liu, Y. H. (2005). *La entonación del español hablado por taiwaneses*. Biblioteca Phonica. www.ub.es/lfa

Liu, C., & Rodriguez, A. (2012). Categorical perception of intonation contrasts: Effects of listeners’ language background. *The Journal of the Acoustical Society of America*, *131*(6), EL427–EL433. https://doi.org/10.1121/1.4710836

Meng, Y., Zhang, J., Liu, S., & Wu, C. (2020). Influence of different acoustic cues in L1 lexical tone on the perception of L2 lexical stress using principal component analysis: an ERP study. *Experimental Brain Research*, *238*(6), 1489–1498. https://doi.org/10.1007/s00221-020-05823-w

Moradi, F., & Shahrokhi, M. (2014). The effect of listening to music on Iranian children’s segmental and suprasegmental pronunciation. *English Language Teaching*, *7*(6), 128–142. https://doi.org/10.5539/elt.v7n6p128

Navarro-Tomás, T. (1974). *Manual de entonación española* (Vol. 175). Ediciones Guadarrama.

Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica*, *41*(1), 1–16.

Ortega-Llebaria, M., Nemogá, M., & Presson, N. (2017). Long-term experience with a tonal language shapes the perception of intonation in English words: How Chinese-English bilinguals perceive “Rose?” vs. “Rose.” *Bilingualism*, *20*(2), 367–383. https://doi.org/10.1017/S1366728915000723

Ortega-Llebaria, M., & Wu, Z. (2021). Chinese-English speakers’ perception of pitch in their non-tonal language: Reinterpreting English as a tonal-like language. *Language and Speech*, *64*(2), 467–487. https://doi.org/10.1177/0023830919894606

Shang, P. (2022a). Cross-linguistic comparison of the pitch and temporal profiles between L1 and Chinese L2 speakers of Spanish. *Loquens*, *9*(1-2), 1–11. https://doi.org/10.3989/loquens.2022.e086

Shang, P. (2022b). *Adquisición de la entonación de las oraciones interrogativas del español por parte de los sinohablantes: patrones entonativos y propiedades fonético-acústicas* [Unpublished doctoral dissertation]. Universitat de Barcelona.

Shang, P., & Elvira-García, W. (2022). Second language acquisition of Spanish prosody by Chinese speakers: Nuclear contours and pitch characteristics. *Vigo International Journal of Applied Linguistics*, (19), 129–176. https://doi.org/10.35869/vial.v0i19.3762

Shang, P., Elvira-García, W., & Li, X. (2022). Cue weighting differences in perception of Spanish sentence types between native listeners of Chinese and Spanish. *Proceedings of the 11th International Conference on Speech Prosody* (pp. 644–648). Lisbon, Portugal. doi: 10.21437/SpeechProsody.2022-131

Shang, P., Elvira-García, W., & Roseano, P. (2021). La modalidad interrogativa en español y en chino: un enfoque funcionalista. *Círculo de Lingüística Aplicada a la Comunicación*, *88*, 235–254. https://doi.org/10.5209/clac.78313

Shang, P., Roseano, P., & Elvira-García, W. (2024). Dynamic multi-cue weighting in the perception of Spanish intonation: Differences between tonal and non-tonal language listeners. *Journal of Phonetics*, *102*. http://dx.doi.org/10.1016/j.wocn.2023.101294

So, C. K., & Best, C. T. (2010). Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and Speech*, *53*(2), 273–293. https://doi.org/10.1177/0023830909357156

R Core Team. (2020). *R: A language and environment for statistical computing* (Version 4.3.1) [Computer software]. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org

Trimble, J. C. (2013). Perceiving intonational cues in a foreign language: Perception of sentence type in two dialects of Spanish. In C. Howe, S. E. Blackwell, & M. L. Quesada (Eds.), *Proceedings of the 15th Hispanic Linguistics Symposium* (pp. 78–92). University of Georgia, USA.

Venables W. N., & Ripley B. D. (2002). *Modern applied statistics with S* (4th ed.). Springer. https://link.springer.com/book/10.1007/978-0-387-21706-2

Wang, M., Zi, G., Xiong, W., & Lin, M. (2013). Pitch declination of Chinese spontaneous speech. *Journal of Chinese Information Processing*, *27*(6), 128–133.

Wiener, S., & Goss, S. (2018). Perceptual assimilation of non-native prosodic cues: Cross-linguistic effects of lexical F0 learning. *Proceedings of the 9th International Conference on Speech Prosody* (pp. 947–951). Poznań, Poland.

Yuan, J. (2011). Perception of intonation in Mandarin Chinese. *The Journal of the Acoustical Society of America*, *130*(6), 4063–4069. https://doi.org/10.1121/1.3651818

Yuan, C., González-Fuente, S., Baills, F., & Prieto, P. (2019). Observing pitch gestures favors the learning of Spanish intonation by Mandarin speakers. *Studies in Second Language Acquisition*, *41*(1), 5–32. https://doi.org/10.1017/S0272263117000316

Zarate-Sandez, G. A., Morales-Front, A., & Sanz, C. (2015). *Perception and production of intonation among English-Spanish bilingual speakers at different proficiency levels.* [Doctoral dissertation, Georgetown University]. Georgetown University Institutional Repository. http://hdl.handle.net/10822/761510

Zhang, S., Li, K., Lo, W.-K., & Meng, H. (2010). Perception of English suprasegmental features by non-native Chinese learners. *Proceedings of the 5th International Conference on Speech Prosody* (pp. 1–4). Chicago, USA.

7. Appendix

Supplemental materials for this study can be found online at https://osf.io/jcbhs/