Loquens 9(1-2)
December 2022, e086
eISSN 2386-2637
https://doi.org/10.3989/loquens.2022.e086

Cross-Linguistic Comparison of the Pitch and Temporal Profiles between L1 and Chinese L2 Speakers of Spanish

Peizhu Shang

Beijing Institute of Technology
shangpeizhu@bit.edu.cn ORCID: https://orcid.org/0000-0001-5407-4476

Recibido: 11/10/2021; Aceptado: 14/11/2021; Publicado en línea: 09/06/2023

Citation / Cómo citar este artículo: Shang, Peizhu (2022). Cross-Linguistic Comparison of the Pitch and Temporal Profiles between L1 and Chinese L2 Speakers of Spanish. *Loquens, 9*(1-2), e086. https://doi.org/10.3989/loquens.2021.e086

ABSTRACT: Cross-linguistic studies between intonational languages suggest that there is a universal trend during the L2 learning process regarding pitch and temporal characteristics. We extend these hypotheses to Chinese learners of Peninsular Spanish—a new pairing of tone and non-tone languages. Using six pitch and temporal metrics, we examine how Chinese learners' pitch and temporal profiles deviated from those of L1 native speakers and explore the factors that may contribute to L2 speech deviations. The Discourse Completion Task was conducted to elicit five question types produced by 37 participants, who were divided into three language groups. Consistent with previous literature, our study shows that Chinese L2 learners had a compression of pitch span (at both the utterance and syllable levels) and pitch variability, as well as a strong reduction of pitch change rate, speech rate, and articulation rate compared to L1 Spanish speakers. Most pitch and temporal deviations in L2 Spanish intonation are closely linked to psychological and cognitive attributes rather than being determined by physiological factors or L1 tonal transfer. Moreover, the lack of prosodic knowledge of the target intonation patterns concerning the different question types may also hinder L2 learners from approaching a native-like pitch and temporal profile.

Keywords: Pitch span, pitch variability, temporal features, Chinese speakers of L2 Spanish.

RESUMEN: *Comparación interlingüística de perfiles tonales y temporales entre hablantes nativos de español y aprendices sinohablantes***.-** Algunos estudios interlingüísticos entre lenguas entonativas sugieren que puede existir una tendencia universal durante el proceso de aprendizaje de la L2 con respecto a las características tonales y temporales. Ex-tendemos estas hipótesis a los aprendices chinos de español peninsular —una nueva combinación lingüística entre lenguas tonales y entonativas. Usando seis métricas tonales y temporales, pretendemos examinar cómo los aprendi-ces chinos se desvían de los hablantes nativos en los perfiles tonales y temporales, y explorar los factores que con-tribuyen a las desviaciones en el habla de la L2. Se ha realizado la Tarea de Finalización del Discurso para elicitar cinco tipos de preguntas producidas por los 37 participantes divididos en tres grupos lingüísticos. En línea con la bibliografía anterior, nuestro estudio muestra que los aprendices chinos presentaban una compresión de rango tonal (tanto a nivel oracional como a nivel silábico) y variación tonal, así como una reducción significativa en la tasa del cambio tonal, la velocidad del habla y la tasa de articulación en comparación con los hablantes nativos de español. La mayoría de las desviaciones tonales y temporales en la entonación de la L2 están estrechamente relacionadas con atributos psicológicos y cognitivos más que con factores fisiológicos o con la transferencia tonal de la L1. Además, la falta de conocimiento prosódico de los patrones entonativos relativos a los diferentes tipos de preguntas en la lengua meta también impide que los aprendices de L2 se asimilen a un perfil tonal y temporal similar a los nativos.

Palabras clave: Rango tonal, variación tonal, características temporales, español como L2 producido por sinohablantes.

1. INTRODUCTION

In the last decades, while there has been a growing body of work on the acquisition of non-native Spanish segments (i.e., Chen, 2007; Cobb & Simonet, 2015; Liu, 2019; Morrison, 2003), stress (i.e., Chen, 2007b; Cortés Moreno, 2005; Kim, 2015; Kimura, Sensui, & Takasawa, 2015), prominence (i.e., Kim, 2016; Van Maastricht, Krahmer, & Swerts, 2016), and intonation contours (i.e., Gabriel & Kireva, 2014; Henriksen, Geeslin, & Willis, 2010; Silva & Barbosa, 2017; Trimble, 2013; Yuan et al., 2019), little is known about the acoustic-phonetic realization of pitch and temporal patterns in L2 Spanish, particularly in environments of language contact between tone and non-tone languages such as Chinese and Spanish. Therefore, the goal of the present study is to fill in the gap by examining cross-linguistic differences of pitch and temporal profiles between first- (L1) and second-language (L2) speakers of Peninsular Spanish.

Pitch profiles consist of the oscillations of fundamental frequency (F0) and are claimed to have quasi-universal and language-specific characteristics in human communication (Chen, Gussenhoven, & Rietveld, 2004; Gussenhoven & Chen, 2000). The generalizability in the use of pitch to convey certain paralinguistic meanings is often explained with biologically determined codes. For example, the frequency code proposes that high pitch is related to a small larynx and often serves as a marker of uncertainty, whilst low pitch is associated with a larger organ of production and is used to signal assertiveness (Gussenhoven, 2002; Ohala, 1983). However, despite this commonality, it is broadly recognized that language communities differ from each other in the specific phonetic implementation of pitch patterns, such as register and range. For instance, by combining

the linguistic and the long-term distributional (LTD) measures, Mennen et al. (2012) found that English female speakers had a significantly higher F0 register and a larger F0 span than their German counterparts. Similar cross-linguistic differences in pitch profiles have also been observed for Polish vs. English (Majewski et al., 1972), Russian vs. German (Nebert, 2013), Mandarin vs. English (Keating & Kuo, 2012), Mandarin vs. Japanese (Shi et al., 2014), Slavic and Germanic languages (Andreeva et al., 2014), and many others (see Mennen et al., 2012 and Ordin & Ineke Mennen., 2017 for a review). Apart from the influence of the L1 prosodic system and some physiological factors such as vocal tract length, gender, and age, the language-specific pitch properties are possibly more closely linked to some social-cultural attributes. Unmistakable evidence for this is that Japanese speakers, particularly women, have a higher F0 register and F0 span than native speakers of Chinese (Shi et al., 2014), Dutch (Van Bezooijen, 1995), American English, and Spanish (Hanley et al., 1966). The preference for high pitches shown by Japanese women is explained in the context of their relative powerlessness in social status and the gender roles they are expected to play according to cultural conventions.

Furthermore, since the speech of a foreign language often entails some degree of interaction, the cross-language differences between the first and the second language can be expected to impact the target speech patterns. Studies have shown that most L2 segmental and suprasegmental errors could be attributed to a prosodic transfer from the L1 system into the phonetic and phonological knowledge of the L2 (Graham & Post, 2018; Ineke Mennen, 2015). However, importantly, several studies have found that some deviated use of pitch is common in L2 speech, revealing itself as a consistent development trajectory during the L2 speech-learning process. For example, the results in previous literature (i.e., Busà & Urbani, 2011; Chen, 1972; Mennen, Schaeffler, & Dickie, 2014; Shi et al., 2014; Ullakonoja, 2007; Yuan et al., 2018) suggest that foreign speakers, regardless of their L1–L2 backgrounds, are often characterized by a narrower F0 range and less variable pitch when producing the L2 speech on the utterance level. In contrast, on the phonemic level, Chinese L2 speakers were reported to have a wider pitch span and smaller F0 fluctuations than native English speakers, mostly due to the negative attachment of L1 lexical tones to stressed syllables in the L2 (Ding et al., 2016; J. Yuan et al., 2018).

The difficulty of accurately implementing the target pitch profiles has been mainly correlated with the L2 learners' lack of confidence and insecurity when speaking a foreign language (Ding et al., 2016; Shi et al., 2014; Yuan et al., 2018), and not merely due to the language specificities and the different socio-cultural identities. Another plausible factor that may constrain the pitch variance is the learners' increased cognitive efforts in producing segments and stress (Zimmerer et al., 2014). Nevertheless, fortunately, studies showed that, with the aid of speech technology or with developing their proficiency in L2, learners were able to fine-tune the production of the L2 pitch and finally approach native-like pitch patterns (Hincks & Edlund, 2009; Ullakonoja, 2007).

On the other hand, L2 speech is also found to be characterized by a decrease in oral fluency (Peters, 2019). The differences in fluency between the L1 and the L2 are frequently measured by various temporal metrics. For example, Ding et al. (2016) showed that, in comparison with native English speakers, Chinese learners tend to have a lower speech rate and articulation rate in their L2 English. Lee and Sidtis (2017) and Peters (2019) made similar observations. The decrease in speech fluency in the non-native language has been explained with reference to the same psychological and cognitive factors as L2 pitch compression—cautiousness and increased cognitive efforts when speaking a foreign language. However, unlike the two variables of speech rate and articulation rate, the temporal assumption of pitch change rate is controversial, especially when it is examined in a stress language such as English compared to a tone language like Chinese. For instance, Yuan et al. (2018) reported a faster pitch change rate for L1 English speakers than for L2 Chinese learners, while in Ding et al. (2016), there was no significant difference between the two language groups with regards to the speed of pitch changes.

Despite the large body of cross-linguistic analyses of pitch and temporal differences, it is somewhat difficult to compare the results of these findings. This is partly because the F0 estimation methods and the fluency measures used for evaluating the pitch and temporal properties differed across studies. Another aspect is that the distinct discourse conditions designed to elicit the speech may also cause inconsistent results. For instance, Yuan and Liberman (2014) reported that Chinese native speakers have a wider pitch range and greater F0 fluctuations in broadcast news speech than native English speakers. However, regarding prose passages (Keating & Kuo, 2012), there was no significant difference in pitch range on the utterance level between Chinese and English speech.

Given the inconsistency of prior results and the typological differences between Chinese and Spanish, it is of great importance to examine the pitch and temporal characteristics in the CH-ES language pair, which has received little attention in the prosodic field to date. Of particular interest to us is to investigate (1) whether the pitch and temporal profiles produced by L2 Chinese learners are highly dependent on their L1 properties or if they support the L2 general trend hypothesis, (2) whether speakers' pitch and temporal implementations are influenced by the gender and the level of proficiency in Spanish, and finally (3) whether the production of L2 pitch and temporal features reflects different levels of difficulty depending on question type and stress position. For these purposes, we extend the previous studies by accounting for proficiency level, gender, question type, and stress position, which allows us to examine the interaction between proficiency and other fixed factors concerning various pitch and temporal metrics.

2. METHODOLOGY

2.1. Participants

The participants of this study included: 5 female native speakers of Peninsular Spanish and 32 learners of Spanish (26 females and 6 males) whose first language is Mandarin Chinese. The ages of Chinese learners ranged from 21 to 31 (mean age: 24.09; *SD* = 2.53), while those of L1 Spanish speakers ranged from 18 to 24, with a mean age of 23.2 years (*SD* = 4.87). All subjects were divided into three language groups according to their proficiency level in Spanish: intermediate (B1-B2 level), advanced (C1-C2 level), and native. The Spanish proficiency of most Chinese speakers was judged using the information from their most recent official language qualification DELE (Diploma of Spanish as a Foreign Language). Chinese learners who did not have this certificate (approximately 15%) were asked to self-evaluate their L2 proficiency based on the Spanish language courses they had completed. The criteria for the six levels of European language proficiency were explained to those participants to help them to reach a reliable self-assessment.

Although the age of acquisition and the length of exposure to the target language are reported to influence L2 speech (Cadierno et al., 2020; Kharkhurin, 2008; Pfenninger & Singleton, 2016), we did not control for these variables, as this would have significantly reduced the number of L2 Chinese participants. However, most of the Chinese learners in this study acquired Spanish in adulthood (mean age: 18.81; *SD* = 2.08). Only one subject reported starting to learn Spanish at 12 years of age. All the Chinese participants were in an immersion situation at the time of recording. Although the length of their stay in Spain had varied, the average exposure time of L2 advanced learners (mean length: 22.80 months; *SD* = 18.02) was generally longer than that of L2 intermediate speakers (mean length: 19.13 months; SD = 9.51).

2.2. Task and materials

The corpus was elicited by utilizing the DCT (Discourse Completion Task) technique (Billmyer & Varghese, 2000; Félix-Brasdefer, 2010). Specifically, we designed 15 brief dialogues structured as situational contexts to elicit five question types with different functional meanings in Spanish, namely, information-seeking *yes-no* question ('YN'), information-seeking *wh*-question ('WH'), *disjunctive* question ('DJ'), confirmation-seeking *yes-no* question ('CYN'), and confirmation-seeking *tag* question ('TAG'). The conversational interaction was initiated by an interlocutor with whom the participant was familiar so that politeness-related effects (e.g., power, and social distance) could be minimized (Borràs-Comes, Sichel-Bazin, & Prieto, 2015; Roseano et al., 2015). A sample context for eliciting the disjunctive question is as follows:

- Interlocutor: *Has invitado a un buen amigo a tu piso para una cena. Después de acabar los platos principales, le preguntas si quiere tarta o helado de postre.* (You have invited a good friend to your apartment for dinner. After finishing the main courses, you ask her if she wants cake or ice cream for dessert.)
- Participant: *¿Quieres tarta o helado?* (Do you want cake or ice cream?)

Each of the five question types varied in the nuclear stress position (two positions: penultimate syllable stress—paroxytone; final syllable stress—oxytone). To facilitate L2 speakers' comprehension during the task, all test items consisted of words with high frequency for L1 and L2 Spanish speakers (Tanaka & Terada, 2011).

The recordings took place in a soundproof room with a head-mounted microphone. Speech files were digitalized at a sampling rate of 44.1 kHz and with a quantization precision of 16 bits. Each utterance was saved separately and annotated to a *TextGrid* object in *Praat* (Boersma & Weenink, 2020).

2.3. Data extraction

For the purposes of this paper, two types of measurements were conducted: (a) pitch and (b) temporal measures. In order to extract the pitch information from the utterances, firstly, the ESPS algorithm (‘get F0’) (Talkin, 1995) was automatically conducted in *Praat* with the pitch floor and ceiling set to 70 Hz and 600 Hz, respectively. A time step of 10 ms was used for the computation of F0. After the automatic extraction, the raw F0 data were corrected manually, unvoicing those pitch points with octave jumps or measurement errors, such as false voicing in silent fragments, creaky voice, and laryngealization. The linear results in Hz were then transformed into the near-logarithmic scale (ERB-rate), which is one of the best psycho-acoustic measures for modeling the intonational equivalence between men and women, and for capturing the F0 differences across languages (Nolan, 2003).

In specific, pitch characteristics in this study were evaluated by means of the three F0 variables: (1) 80% pitch span on the utterance level (the 90th and 10th percentile span), (2) absolute span on the syllable level (the 100th percentile span), and (3) pitch dynamism quotient (abbreviated as PDQ). The PDQ metric was included as a normalization of the F0 variation data since it can minimize the effects caused by gender and different group size. The PDQ value gives an account of the pitch variability in the utterance, and it is calculated by dividing the standard deviation by the F0 mean. In general, the previous literature indicates that the higher the PDQ, the more variable the speech (Shi, Zhang, & Xie, 2014; Wang & Qian, 2018; Zimmerer et al., 2014).

Further, considering the temporal traits, three variables were examined between L1 and L2 speech: (1) pitch change rate (the average of the absolute pitch differences in every 10-ms interval), (2) speech rate (number of syllables / total duration of the utterance), and (3) articulation rate (number of syllables / (total duration–internal pauses). The minimum pause length calculated for fluency judgments was set to 0.05 s instead of the larger values of 0.25 s adopted in the study of Peters (2019). The underlying reason is that the speech materials used in our experiment were single utterances with an average syllable number of 5.8—unlike the passages in Peters (2019) that frequently required the use of long pauses as a linguistic cue for narrative segmentation (Oliveira, 2002).

2.4. Statistical analysis

The data analysis was conducted in the R environment (R Core Team, 2020). A linear mixed-effects analysis was carried out using the *lmerTest* package for R (Kuznetsova et al., 2017). The six pitch and temporal parameters (80% span on the utterance level, PDQ, 100% span on the syllable level, pitch change rate, speech rate, and articulation rate) were entered into the model successively as dependent variables, with *Proficiency Level* in Spanish (intermediate < advanced < native), *Gender* (female vs. male), *Question Type* (i.e., YN, WH, DJ, CYN, and TAG), *Stress Type* (Oxytone vs. Paroxytone), and their possible interactions as fixed effects. Participants were included as random effects with all possible random intercepts. The significance of the main effects was tested using the *ANOVA* function. *P*-values were fitted by eliminating the non-significant effects of the initial model and calculated with Satterthwaite’s method (Kuznetsova et al., 2017). The post-hoc analysis was performed using the single-step function of the *multcomp* package (Hothorn et al., 2016) supported by the *emmeans* algorithm (Lenth et al., 2019).

3. RESULTS

The following two sections present the results of the three pitch variables measured on the utterance (80 % F0 span, and PDQ) and syllable level (100 % F0 span), and the results of the three temporal parameters (pitch change rate, speech rate, and articulation rate).

3.1. Pitch results

First, we considered the differences in the use of pitch across the three language groups. The analysis of variance indicated that *Proficiency Level* was not a significant factor for the three pitch variables (see Table 1). However, Figures 1, 2, and 3 indicate that Chinese intermediate (hereafter CI) and advanced learners (hereafter CA) tend to produce a less variable pitch and narrower span on the utterance and syllable levels compared to L1 Spanish speakers (hereafter SN). These findings generally are consistent with previous studies that reported a reduced pitch for non-native speakers (Busà & Urbani, 2011; Mennen, Schaeffler, & Docherty, 2007; Shi et al., 2014; Yuan et al., 2018; Zimmerer et al., 2014), suggesting that there may be a universal trend of pitch range compression in L2 speech. Additionally, the results in Figures 1, 2, and 3 indicated that, in comparison with the lower proficiency

Table 1: Effects (*F*-values) of Proficiency level, Question type, Gender, Stress position, and their interactions on the three pitch variables (‘***’ $p < 0.001$; ‘**’ $p < 0.01$; ‘*’ $p < 0.05$; ‘.’ $p < 0.1$).

	80 % utterance span	PDQ	100 % syllable span
Proficiency	2.99.	2.80.	2.53.
QuestionType	10.53***	8.99***	22.26***
Gender	0.00	8.76**	1.33
StressType	3.12.	4.00*	0.42
Proficiency*QuestionType	9.98***	8.42***	8.58***
Proficiency* StressType	3.54*	0.22	0.37

Table 2: Pairwise comparisons between the three language groups regarding the three pitch variables.

	80 % utterance span	100 % syllable span	PDQ
CI-CA	$t = -1.802$, $p = 0.179$	$t = -1.783$, $p = 0.185$	$t = -1.791$, $p = 0.182$
SN-CA	$t = 1.029$, $p = 0.559$	$t = 0.766$, $p = 0.723$	$t = 0.932$, $p = 0.620$
SN-CI	$t = 2.289$, $p = 0.067$	$t = 2.029$, $p = 0.116$	$t = 2.190$, $p = 0.083$

group (i.e., the CI group), highly proficient learners of the CA group were closer to SN speakers in the implementation of the F0 pitch, although this trend was not strong enough to be statistically significant (see Table 2).

Next, as with the *Question Type* factor, it is apparent in Table 1 that there is a significant main effect on the three pitch variables. In contrast, the factors *Gender* and *Stress Type* were found to be significant only for the variable of PDQ. In particular, our results indicated that female speakers (mean PDQ: 0.175) had significantly more F0 variability than males (mean PDQ: 0.127) in speech [$t(70) = 2.14$, $p < 0.05$). We also observed a significant effect of *Stress Type* on the variable of PDQ. Specifically, it is noteworthy in Figure 2 (see the right panel) that participants of the three language groups consistently had a more variable pitch in questions with a paroxytone than those with an oxytone in the final word.

As with the 80% utterance span, Figure 1 shows that the two Chinese groups had a wider pitch span in questions ending with a paroxytone word, but this tendency was statistically significant only for the CA group [$t(539) = 3.07$, $p < 0.01$]. Regarding the SN group, we did not find a statistically significant difference in realizing the pitch between the two stress types [$t(539) = 0.04$, $p = 0.76$], although SN speakers were more likely to compress the F0 span in questions ending with a paroxytone word (see the right panel of Figure 1). The pitch performance exhibited by the CI and CA groups may be because the paroxytone is the most frequent and unmarked stress pattern in Spanish and, therefore, the most familiar one for L2 speakers (Defior & Serrano, 2017; Roca, 2019). This means that Chinese learners may experience the least cognitive difficulties when producing such stressed words in Spanish, which allows more planning time to fine-tune the corresponding pitch profiles in a native-like way. In contrast, it is unclear why SN speakers had an opposite trend for implementing the F0 span between the two stress types. Since we only had five Spanish subjects in this work, future investigations with a larger sample size are needed to validate this finding.

The results of the linear mixed model also revealed a strong interaction effect between *Proficiency Level* and *Question Type* on the three pitch variables (see Table 1). The post-hoc analysis indicated that the pitch performance of CI and CA learners was highly dependent on the question type in which they were engaged. More precisely, we found that, in comparison with the SN group, the CI and CA group had a particularly narrower span and less pitch variability in DJ [e.g. 80% span: CI-SN: $t(2) = -4.04$, $p < 0.001$; CA-SN: $t(2) = -3.79$, $p < 0.001$] and YN questions [e.g. PDQ: CI-SN: ($t(2) = -3.35$, $p < 0.01$); CA-SN: ($t(2) = -2.47$, $p < 0.05$)]. By contrast, in WH questions, it is noteworthy that the two Chinese groups had a higher PDQ and a wider pitch span on both utterance and syllable levels than the SN group (see Figures 1, 2, and 3). This finding can be explained by the overproduction of WH questions by Chinese learners. Specifically, we notice that some L2 learners, irrespective of their level of proficiency, tend to produce a high-rising nuclear pitch accent or a final rising boundary tone in WH questions. Although the final rising contour can also be used in WH questions, it is not frequently found in the L1 native speech (i.e., all

Figure 1: 80 % pitch span on the utterance level of the three language groups depending on Question Type (left) and Stress Type (right). Error bars indicate ± 1SE.

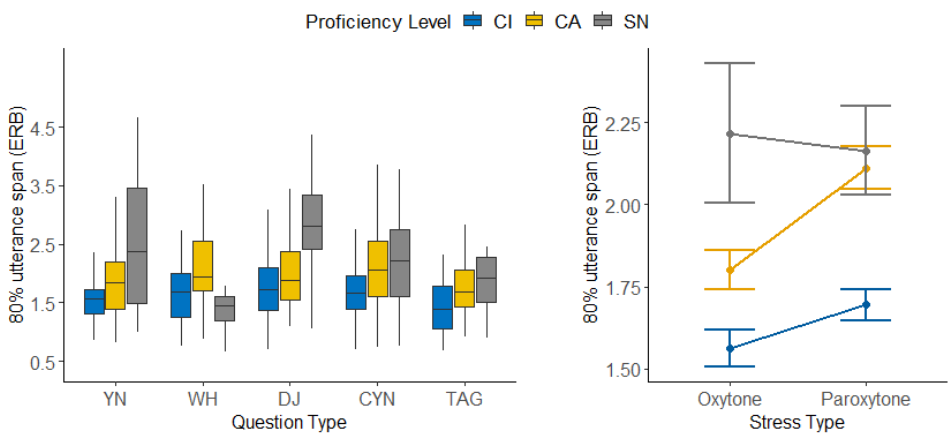


Figure 2: Mean PDQ of the three language groups depending on Question Type (left) and Stress Type (right). Error bars indicate ± 1SE.

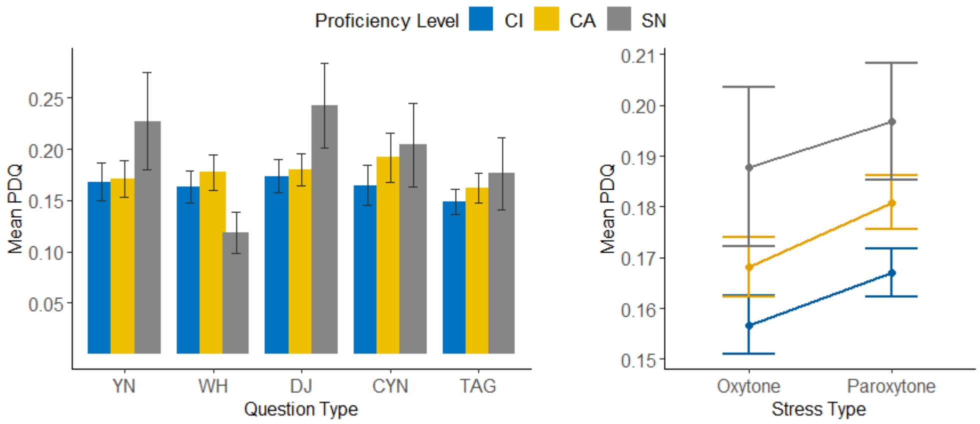
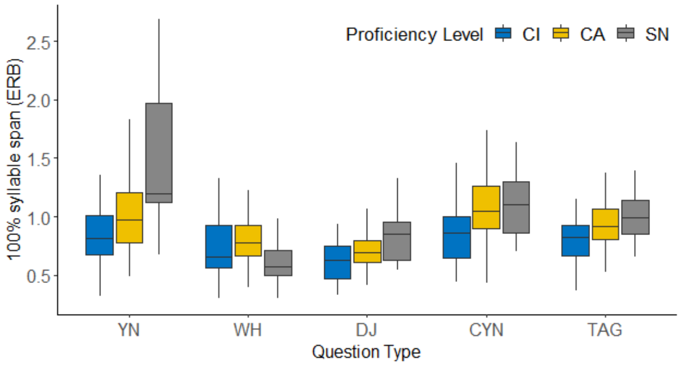


Figure 3: 100 % pitch span on the syllable level depending on Proficiency Level and Question Type.



the SN speakers in our study produced the WH questions with a final-falling pitch movement) since the interrogative particles in Spanish (e.g., *qué, dónde, quién, cuál*) are clear enough for signaling this type of question.

3.2. Temporal results

The main effects of the linear mixed models fitted for the three temporal variables are shown in Table 3. For ease of exposition, we discuss these results by referring to Figures 4 and 5, which display the specific temporal values produced by the three language groups in the five question types. First, considering individual effects, the output in Table 3 revealed that there was a significant main effect of *Proficiency* and *Question Type* on the outcome variables of pitch change rate, speech rate, and articulation rate. By contrast, *Stress Type* and *Gender* were insignificant factors for the three temporal variables. Moreover, the pairwise comparisons of *Proficiency Level* showed that, in comparison with the SN group, the two Chinese groups had a significantly lower pitch change rate [CI-SN: $t(2) = -4.71$, $p < 0.001$; CA-SN: $t(2) = -3.75$, $p < 0.01$], speech rate [CI-SN: $t(2) = -5.71$, $p < 0.001$; CA-SN: $t(2) = -5.62$, $p < 0.001$], and articulation rate [CI-SN: $t(2) = -5.58$, $p < 0.001$; CA-SN: $t(2) = -5.44$, $p < 0.001$] in their speech. These findings corroborate previous studies that reported a reduced oral fluency for L2 speakers in the non-native language (Ding et al., 2016; Peters, 2019). Nevertheless, unlike our previous findings—which showed that high-proficiency Chinese learners achieved a target-like pitch performance—(see Section 3.1), we did not observe any significant improvement in speech rate and articulation rate between the CI and CA groups.

Further, the results in Table 3 indicated a strong interaction between *Proficiency* and *Question Type* on the three temporal variables. Particularly, as shown in Figure 5, SN speakers had higher values of pitch change rate than the CI and CA learners in all questions except for WH questions. As discussed above, the faster pitch change in L2 WH questions may be attributed to the fact that most Chinese learners excessively varied their F0 contours by producing either a high pitch accent or a final rising boundary in the nuclear position. In addition, although each question type was realized with a specific temporal value, the two Chinese groups were consistently lower than the SN speakers regarding the speech and articulation rates (see Figure 5). Finally, it is interesting that the results of speech rate and articulation rate were similar in this work. This is perhaps because the speech stimuli used in this work consisted of short utterances produced with low frequency and short pauses.

4. DISCUSSION

The aim of the present study was to investigate the L2 production of Spanish questions by Chinese speakers with regards to pitch and temporal characteristics and to explore the factors that may contribute to the pitch and temporal deviations in L2 speech. Six pitch and temporal metrics of L1 and L2 Spanish speakers were examined and compared using a linear mixed-effects analysis. The findings of our study are discussed below.

First, our results confirm that there are indeed some cross-linguistic differences between Spanish L1 and L2 regarding pitch performance. The evidence in support of this is that the L2 Spanish in this study was produced with a narrower span (on both utterance and syllable levels) and less variable pitch than that of L1 native speakers. This supports previous studies that reported a pitch range compression effect for L2 speakers with typologically different L1 backgrounds (e.g., Busà & Urbani, 2011; Liu, 2005; Mennen et al., 2007, 2012, 2014; Peters, 2019; Shi et al., 2014; Ullakonoja, 2007; Urbani, 2012; Yuan et al., 2018; Zimmerer et al., 2014). The consistency of the findings for L2 pitch and temporal production suggests that non-native learners may have universal developmental pathways for acquiring specific aspects of L2 speech, independent of the specificity of their L1 system. We cannot

Table 3: Effects (*F*-values) of Proficiency Level, Question type, Stress Type, and their interactions on the three temporal variables ('***' $p < 0.001$; '**' $p < 0.01$; '*' $p < 0.05$; '.' $p < 0.1$).

	Pitch change rate	Speech rate	Articulation rate
Proficiency	11.23***	18.75***	17.75***
QuestionType	14.95***	10.56***	4.35**
Gender	0.03	2.71	3.26.
StressType	0.01	0.22	0.00
Proficiency*QuestionType	11.46***	3.80***	2.64**

Figure 4: Pitch change rate of the three language groups depending on Proficiency Level and Question Type. Error bars indicate ± 1SE.

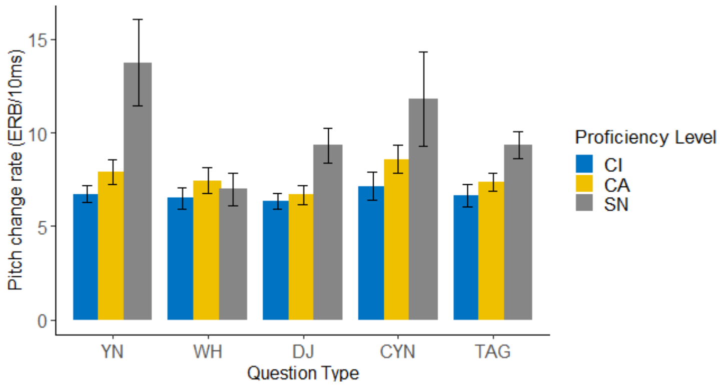
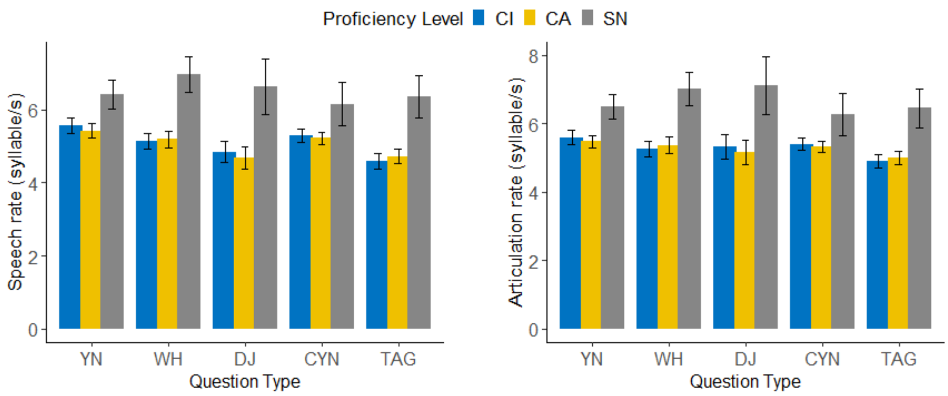


Figure 5: Speech rate (left) and articulation rate (right) of the three language groups depending on Proficiency Level and Question Type. Error bars indicate ± 1SE.



provide a definitive explanation for this quasi-universal effect in L2 speech. However, rather than being shaped by the L1 phonetic system, the compressed pitch patterns in L2 have previously been attributed to the lack of confidence and insecurity of L2 learners when speaking a non-native language (Peters, 2019; Volín, Poesová, & Weingartová, 2015; Yuan et al., 2018). Additionally, the increased cognitive efforts in producing the non-native segmental or suprasegmental features (i.e., vowels and consonants, stress, and prominence) are also plausible factors that may lead to a lower pitch variability in L2 utterances. For instance, Zimmerer et al. (2014) pointed out that L2 learners can frequently overlook the variation of F0 pitch in a native-like way because they are too focused on the correct production of words and stress in the non-native language.

Another noteworthy point in the pitch results is the F0 span at the syllable level. As a typical tone language, Chinese makes use of the F0 information for encoding lexical tone meanings (Yuan, 2011). Therefore, it is expected that Chinese learners would show greater F0 variations on the syllable level because of L1 tonal transfer. However, unlike Ding et al. (2016), we did not find a wider pitch span on the syllable level for Chinese learners of Spanish. This seems to imply that the production of the L2 syllable span was not necessarily affected by the learners' long-term experience with a tone language. The discrepancy between the results could be justified by the distinct language pairs examined in the experiment: In Ding et al. (2016), English was the Chinese learners' L2, whereas in our study, it was Spanish. Future studies regarding the pitch range differences between English and Spanish at the syllable level would help us elucidate whether this is the primary cause of the discordances found. On the other side, based on our observed data, another possible explanation for the reduced syllable span in L2 Spanish might be that Chinese learners were too cautious to vary the pitch due to a lack of intonational skills and language experience, thereby exhibiting a flat F0 contour without many fluctuations until they reached the great F0 changes in the nuclear location. Further investigations of L2 phonetic performance are required to test this hypothesis, considering the position sensitivity of pitch changes in the utterance.

Further, although the factor proficiency statistically failed to reach significance in the three pitch variables, the results seem to suggest that Chinese learners of L2 Spanish can progressively fine-tune their production of F0 values and approach a target-like pitch pattern with increasing proficiency in their L2. Moreover, results of the three pitch parameters revealed a strong interaction between proficiency level and question type, illustrating that the L2 learning of pitch implementation details is susceptible to pragmatically different question types. For instance, we found that Chinese intermediate and advanced learners consistently had a reduced pitch span and lower PDQ in all utterances except for WH questions. As is clear from the above discussion, the opposite performance of Chinese speakers on WH-questions can account for their overproduction of a high pitch accent or a final-rising boundary in the nuclear position. Or, in a more general way, it can be attributed to the fact that learners were unfamiliar with the target intonation contours of WH-questions due to the typological distance between the L1 and the L2. Thus, most would simply assume that Spanish WH-questions are produced with a high pitch in the utterance-final location based on their knowledge of the typical use of the F0 cue.

As with other question types (e.g., the information-seeking *yes-no* question and the disjunctive question), we found that most F0 targets in the utterance-final position could be accurately achieved by Chinese learners, while those in the prenuclear position were deviated and produced with a less variable contour. In this regard, our findings suggest that the compressed pitch in L2, rather than being solely determined by psychological and cognitive factors (i.e., uncertainty, cautiousness, and increased efforts when speaking the L2), is also constrained by the learners' knowledge of the target intonation categories. Overall, the different pitch performance of the L2 speakers in the five question types gives support to previous findings which proposed a scaffolding from the phonological to phonetic dimensions (Cortés Moreno, 2004; Yuan et al., 2019), suggesting that there is a hierarchy of difficulties in implementing the L2 pitch patterns depending on the prosodic similarities and dissimilarities between the first and the target language.

Considering the gender effect, our study revealed that men and women differ significantly only in the variable of PDQ. Congruent with previous works (Ordin & Ineke Mennen., 2017), female speakers in our study varied their F0 contours more frequently than male speakers. The gender differences in pitch variability are more closely linked to the speakers' willingness to express emotions in communication rather than physiological factors. Research has shown that humans express a range of emotions by readily modulating their F0 pitch, and female speakers tend to express most emotions more frequently than males in speech—except for pride and power (Brebner, 2003; Pisanski et al., 2020). In this sense, the greater pitch variance observed in the data of female speakers could be attributed to their greater emotional involvement in speech than male participants. Nevertheless, because the number of male and female speakers differed strongly in this task, this research needs to be replicated with a well-balanced design to consolidate the results presented here.

Further interesting findings related to pitch are that the F0 variation was highly modulated by the stress type, whereby all speakers produced more variable pitch in questions with a final-paroxytone word than in those with a final-oxytone word. Similarly, for the 80 % F0 span, Chinese learners (particularly those of the advanced group) showed a significantly wider pitch span in questions ending with a paroxytone word. We speculate that this could be related to the relative cognitive efforts required to process the two stress types for L2 learners. Since the paroxytone is the most frequent and unmarked stress pattern in Spanish (hence the most familiar one for L2 learners), Chinese speakers may show fewer difficulties when producing it in questions and have more planning time, allowing them to better approach a target-like pitch profile. Although L1 Spanish speakers had a reduced pitch span in sentences with a final-paroxytone word, this effect did not reach statistical significance, and their average pitch span was still higher than that of Chinese learners with such stimuli. So far, we have no clear explanation for the behaviour of Spanish speakers. Since there were only five native subjects in the control group, future investigations with a larger sample size are required to test whether there is a difference of pitch span for L1 Spanish speakers in questions ending with different stress patterns.

Regarding the temporal characteristics, our study revealed significantly lower pitch change rate, speech rate, and articulation rate in L2 Spanish. These results are consistent with previous studies that reported a similar reduction of oral fluency (Ding et al., 2016; Peters, 2019) and slower pitch rises and falls in L2 speech than L1 speech (Yuan et al., 2018). Moreover, it has been noted that although Chinese is a lexical tone language with F0 peaks or valleys in every syllable, the speed of F0 changes is not significantly faster than in stress languages such as English (Xu & Sun, 2002). If this is the case, we speculate that there is no negative transfer of L1 Chinese in terms of the pitch change rate in this study. The lower values of Chinese L2 learners on the three temporal metrics might also be attributed to their increased cognitive efforts in producing the segments or their lack of experience in the target speech.

Additionally, the interaction effects found for the three temporal variables indicate that the proficiency effect was strongly modulated by question type. Whereas the speech rate and articulation rate were lower in all question types for L2, the average pitch change rate showed an exception for the WH questions in which the F0 directions varied more frequently in L2 than in L1. Since there is no indication that the L2 deviation on WH questions was caused by the systematic differences between the two languages, we speculate that the higher values of pitch change rate and F0 span in WH questions reflected overproduction by Chinese speakers due to a lack of target intonational knowledge. Finally, the main effect of proficiency seems to suggest a trend of pitch improvement with learners' increasing L2 proficiency. In particular, our study replicates previous findings (i.e., Ullakonoja, 2007; Yuan et al., 2018; Zimmerer et al., 2014) that highly proficient learners were closer to L1 native speakers in the realization of pitch change rate, pitch span on the utterance and syllable level, and pitch variability. Further, as suggested by neurobehavioral research, the advantage of high-proficiency speakers in the L2 can be attributed to their enhanced ability to use higher-level cognition (i.e., attention) to process non-native speech components (Archila-Suerte et al., 2012, 2015).

5. CONCLUSION

The study presented here was intended to explore the pitch and temporal characteristics of native and Chinese L2 speakers of Spanish. Using six different metrics, we examined the pitch and temporal implementation in five question types of Peninsular Spanish and obtained several important findings regarding the cross-linguistic differences in the speech. First, congruent with previous literature on L2 speech, the results of this study suggest that Chinese speakers of L2 Spanish deviate from L1 native speakers mainly in the compression of pitch span (both on the utterance and syllable levels) and pitch variability, and the strong reduction of pitch change rate, speech rate, and articulation rate. Second, these pitch and temporal deviations in L2 speech are attributed to psychological-cognitive factors and the learners' lack of knowledge and intonation skills in the target language rather than physiological factors or the L1 effect.

From the pedagogical perspective, our findings hold important implications for understanding the cross-linguistic differences between L1 and the speech, underlining the importance of preparing special training methods with varied materials and contexts to reduce learners' foreign accents and improve their phonetic knowledge of the L2. Further research on native Chinese and native Spanish will be conducted to explore more cross-linguistic differences that may account for the L2 speech deviations. It is also interesting to consider how pitch span and pitch variability are realized depending on the syntactic and phonological positions of the phrase and in which locations L2 learners mostly deviate from the L1 native speakers.

6. ACKNOWLEDGEMENTS

The author is extremely grateful to Doctor Wendy Elvira-García for her technical support. An earlier version of this paper was presented at the 4th Phonetics and Phonology in Europe conference (Barcelona, June 21-23, 2021), and I want to thank the audience there for their feedback. This work has been funded by a Ph.D. research grant provided by the China Scholarship Council (No. 201808390059).

7. REFERENCES

Andreeva, B., Demenko, G., Wolska, M., Möbius, B., Zimmerer, F., Jügler, J., Oleskowicz-Popiel, M. & Trouvain, J. (2014). Comparison of pitch range and pitch variation in Slavic and Germanic languages. In *Proceedings to the 7th Speech Prosody Conference* (pp. 776–780). Dublin, Ireland.

Archila-Suerte, P., Zevin, J., Bunta, F., & Hernandez, A. E. (2012). Age of acquisition and proficiency in a second language independently influence the perception of non-native speech. *Bilingualism*, *15*(1), 190.

Archila-Suerte, P., Zevin, J., & Hernandez, A. E. (2015). The effect of age of acquisition, socioeducational status, and proficiency on the neural processing of second language speech sounds. *Brain and Language*, *141*, 35–49.

Billmyer, K., & Varghese, M. (2000). Investigating instrument-based pragmatic variability: Effects of enhancing discourse completion tests. *Applied Linguistics*, *21*(4), 517–552.

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

Borràs-Comes, J., Sichel-Bazin, R., & Prieto, P. (2015). Vocative intonation preferences are sensitive to politeness factors. *Language and Speech*, *58*(1), 68–83.

Brebner, J. (2003). Gender and emotions. *Personality and Individual Differences*, *34*(3), 387–394.

Busà, M. G., & Urbani, M. (2011). A Cross Linguistic Analysis of Pitch Range in English L1 and L2. In *XVII International Congress of Phonetic Sciences* (pp. 380–383). Hong Kong, China.

Cadierno, T., Hansen, M., Lauridsen, J. T., Eskildsen, S. W., Fenyvesi, K., Jensen, S. H., & aus der Wieschen, M. V. (2020). Does younger mean better? Age of onset, learning rate and shortterm L2 proficiency in young Danish learners of English. *Vigo International Journal of Applied Linguistics*, *17*, 57–86.

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

Chen, G. T. (1972). *A comparative study of pitch range of native speakers of Midwestern English and Mandarin Chinese: An acoustic study* [Doctoral dissertation]. University of Wisconsin.

Chen, Y. (2007a). *A comparison of Spanish produced by Chinese L2 learners and native speakers—an acoustic phonetics approach* [Doctoral dissertation]. Department of Philosophy, University of Illinois at Urbana-Champaign.

Chen, Y. (2007b). From tone to accent: the tonal transfer strategy for Chinese L2 learners of Spanish. In *Proceedings of 16th International Congress of Phonetic Sciences* (pp. 6–10). Saarbrücken, Germany.

Cobb, K., & Simonet, M. (2015). Adult second language learning of Spanish vowels. *Hispania*, 47–60.

Cortés Moreno, M. (2004). Análisis acústico de la producción de la entonación española por parte de sinohablantes. *Estudios de Fonética Experimental*, *13*, 80–110.

Cortés Moreno, M. (2005). Análisis experimental del aprendizaje de la acentuación y la entonación españolas por parte de hablantes nativos de chino. *Phonica*, *1*, 1–25.

Defior, S., & Serrano, F. (2017). Learning to Read Spanish. In L. Verhoeven & C. Perfetti (Eds.), *Learning to Read across Languages and Writing Systems* (pp. 243–269). Cambridge University Press.

Ding, H., Hoffmann, R., & Hirst, D. (2016). Prosodic transfer: A comparison study of f0 patterns in L2 english by chinese speakers. In *8th International Conference on Speech Prosody* (pp. 756–760). Boston, US.

Félix-Brasdefer, J. C. (2010). Data collection methods in speech act performance. *Speech Act Performance: Theoretical, Empirical and Methodological Issues*, *26*(41), 69–82.

Gabriel, C., & Kireva, E. (2014). Prosodic transfer in learner and contact varieties: Speech rhythm and intonation of Buenos Aires Spanish and L2 Castilian Spanish produced by Italian native speakers. *Studies in Second Language Acquisition*, *36*(2), 257–281.

Graham, C., & Post, B. (2018). Second language acquisition of intonation: Peak alignment in American English. *Journal of Phonetics*, *66*, 1–14.

Gussenhoven, C. (2002). Intonation and Interpretation : Phonetics and Phonology. In *International Conference on Speech Prosody 2002* (pp. 47–57). Aix-en-Provence, France.

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

Hanley, T. D., Snidecor, J. C., & Ringel, R. L. (1966). Some acoustic differences among languages. *Phonetica*, *14*(2), 97–107.

Henriksen, N. C., Geeslin, K. L., & Willis, E. W. (2010). The development of L2 Spanish intonation during a study abroad immersion program in León, Spain: Global contours and final boundary movements. *Studies in Hispanic and Lusophone Linguistics*, *3*(1), 113–162.

Hincks, R., & Edlund, J. (2009). Using speech technology to promote increased pitch variation in oral presentations. In *International Workshop on Speech and Language Technology in Education* (pp. 1–4). Wroxall, UK.

Hothorn, T., Bretz, F., Westfall, P., Heiberger, R. M., Schuetzenmeister, A., Scheibe, S., & Hothorn, M. T. (2016). Package 'multcomp.' *Simultaneous Inference in General Parametric Models. Project for Statistical Computing*, Vienna, Austria.

Keating, P., & Kuo, G. (2012). Comparison of speaking fundamental frequency in English and Mandarin. *Journal of the Acoustical Society of America*, *132*(2), 1050–1060.

Kharkhurin, A. V. (2008). The effect of linguistic proficiency, age of second language acquisition, and length of exposure to a new cultural environment on bilinguals' divergent thinking. *Bilingualism: Language and Cognition*, *11*(2), 225–243.

Kim, J.-Y. (2015). Perception and production of Spanish lexical stress by Spanish heritage speakers and English L2 learners of Spanish. In *Proceedings of the 6th Conference on Laboratory Approaches to Romance Phonology* (pp. 106–128). Cascadilla, Somerville, MA.

Kim, J. Y. (2016). *The perception and production of prominence in Spanish by heritage speakers and L2 learners* [Doctoral dissertation]. Department of Spanish and Portuguese, University of Illinois at Urbana-Champaign.

Kimura, T., Sensui, H., & Takasawa, M. (2010). A Pilot Study on Perception of Spanish Stress by Japanese Learners of Spanish. In *Second Language Studies: Acquisition, Learning, Education and Technology* (pp. 1–4). Tokyo, Japan.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13).

Lee, B., & Sidtis, D. V. L. (2017). The bilingual voice: Vocal characteristics when speaking two languages across speech tasks. Speech, *Language and Hearing*, *20*(3), 174–185.

Lenth, R., Singmann, H., & Love, J. (2019). Emmeans: Estimated marginal means, aka least-squares means. *R Package Version 1.3.4*.

Liu, Y. H. (2005). *La entonación del español hablado por taiwaneses*. Biblioteca Phonica, 2. www.ub.es/lfa.

Liu, Z. (2019). *Análisis de las obstruyentes en chino y en español como L3: estudio acústico y perceptivo para la categorización de errores* [Tesis doctoral]. Departamento de Filología Española, Universitat Autònoma de Barcelona.

Majewski, W., Hollien, H., & Zalewski, J. (1972). Speaking fundamental frequency of Polish adult males. *Phonetica*, *25*(2), 119–125.

Mennen, I, Schaeffler, F., & Docherty, G. (2007). Pitching it differently: a comparison of the pitch ranges of German and English speakers. In *16th International Congress of Phonetic Sciences* (pp. 1769–1772). Saarbrücken, Germany.

Mennen, Ineke. (2015). Beyond segments: towards an L2 intonation learning theory (LILt). In *Prosody and language in contact* (pp. 171–188). Berlin, Heidelberg: Springer.

Mennen, Ineke, Schaeffler, F., & Dickie, C. (2014). Second language acquisition of pitch range in german learners of english. *Studies in Second Language Acquisition*, *36*(2), 303–329.

Mennen, Ineke, Schaeffler, F., & Docherty, G. (2012). Cross-language differences in fundamental frequency range: A comparison of English and German. *Journal of the Acoustical Society of America*, *131*(3), 2249–2260.

Morrison, G. S. (2003). Perception and production of Spanish vowels by English speakers. In *Proceedings of the 15th international congress of phonetic sciences* (pp. 1533–1536). Barcelona, Spain.

Nebert, A. U. (2013). *Der Tonhöhenumfang der deutschen und russischen Sprechstimme: Vergleichende Untersuchung zur Sprechstimmlage* (Hallesche Schriften zur Sprechwissenschaft und Phonetik 46). Frankfurt Am Main: Lang.

Nolan, F. (2003). Intonational equivalence: an experimental evaluation of pitch scales. In *Proceedings of the 15th international congress of phonetic sciences* (Vol. 771, pp. 2–5). Barcelona, Spain.

Ohala, J. J. (1983). Cross-language use of pitch: an ethological view. *Phonetica*, *40*(1), 1–18.

Oliveira, M. (2002). The role of pause occurrence and pause duration in the signaling of narrative structure. In *International Conference for Natural Language Processing in Portugal* (pp. 43–51). Springer.

Ordin, M., & Ineke Mennen. (2017). Cross-Linguistic Differences in Bilinguals' Fundamental Frequency Ranges. *Journal of Speech, Language, and Hearing Research*, *60*(6), 1493–1506.

Peters, J. (2019). Fluency and speaking fundamental frequency in bilingual speakers of High and Low German. In *Proceedings of the 19th International Congress of Phonetic Sciences* (pp. 1–5). Melbourne, Australia.

Pfenninger, S. E., & Singleton, D. (2016). Age of onset, socio-affect and cross-linguistic influence: a long-term classroom study. Vigo International *Journal of Applied Linguistics*, *13*, 147–179.

Pisanski, K., Raine, J., & Reby, D. (2020). Individual differences in human voice pitch are preserved from speech to screams, roars and pain cries. *Royal Society Open Science*, *7*(2), 191642.

R Core Team (2020). *R: A language and environment for statistical computing* [Computer software manual]. Vienna, Austria.

Reeder, J. T. (1998). English Speakers' Acquisition of Voiceless Stops and Trills in L2 Spanish. *Texas Papers in Foreign Language Education*, *3*(3), 101–118.

Roca, I. (2019). Spanish Word Stress: an updated multidimensional account. In R. Goedemans, J. Heinz & H. van der Hulst (Eds.), *The Study of Word Stress and Accent: Theories, Methods and Data* (pp. 256–292). Cambridge University Press.

Roseano, P., Fernández Planas, A. M., Elvira-García, W., Massó, R. C., & Celdrán, E. M. (2015). La entonación de las preguntas parciales en catalán. *Revista Española de Lingüística Aplicada*, *28*(2), 511–554.

Shi, S., Zhang, J., & Xie, Y. (2014). Cross-language comparison of F0 range in speakers of native Chinese, native Japanese and Chinese L2 of Japanese: Preliminary results of a corpus-based analysis. In *Proceedings of the 9th International Symposium on Chinese Spoken Language Processing* (pp. 241–244). Singapore.

Silva, C. C., & Barbosa, P. A. (2017). The contribution of prosody to foreign accent: A study of Spanish as a foreign language. *Loquens*, *4*(2), e041–e041.

Talkin, D. (1995). A robust algorithm for pitch tracking (RAPT). In W. B. Klejin & K. K. Paliwal (Eds.), *Speech coding and synthesis* (pp. 495–518). Elsevier Science B.V.

Tanaka-Ishii, K., & Terada, H. (2011). Word familiarity and frequency. *Studia Linguistica*, *65*(1), 96–116.

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.), *15th Hispanic Linguistics Symposioum* (pp. 78–92). Athens, USA.

Ullakonoja, R. (2007). Comparison of pitch range in Finnish (L1) and Russian (L2). In *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 1701–1704). Saarbrücken, Germany.

Urbani, M. (2012). Pitch range in L1/L2 English. An analysis of f0 using LTD and linguistic measures. In M. G. Busà & S. Antonio (Eds.), *Methodological perspectives on L2 prosody: Papers from ML2P 2012* (pp. 79–83).

Van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech*, *38*(3), 253–265.

Van Maastricht, L., Krahmer, E., & Swerts, M. (2016). Prominence Patterns in a Second Language: Intonational Transfer From Dutch to Spanish and Vice Versa. *Language Learning*, *66*(1), 124–158.

Volín, J., Poesová, K., & Weingartová, L. (2015). Speech melody properties in English, Czech and Czech English: Reference and interference. *Research in Language*, *13*(1), 107–123.

Xu, Y., & Sun, X. (2002). Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America*, *111*(3), 1399–1413.

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.

Yuan, J. (2011). Perception of intonation in Mandarin Chinese. *Journal of the Acoustical Society of America*, *130*(6), 4063–4069.

Yuan, J., Dong, Q., Wu, F., Luan, H., Yang, X., Lin, H., & Liu, Y. (2018). Pitch characteristics of L2 English speech by Chinese speakers: A large-scale study. In *Proceedings of the Annual Conference of the International Speech Communication Association* (pp. 2593–2597). Hyderabad.

Yuan, J., & Liberman, M. (2014). F0 declination in English and Mandarin broadcast news speech. *Speech Communication*, *65*, 67–74.

Zimmerer, F., Jügler, J., Andreeva, B., Möbius, B., & Trouvain, J. (2014). Too cautious to vary more? A comparison of pitch variation in native and non-native productions of French and German speakers. In Ni. Campbell, D. Gibbon, & D. Hirst (Eds.), *Proceedings to the 7th Speech Prosody Conference* (pp. 1037–1041). Dublin, Ireland.