Second language acquisition of Spanish prosody by Chinese speakers: Nuclear contours and pitch characteristics

Peizhu Shang
Department of Catalan Philology and General Linguistics
University of Barcelona, Spain
pshangsh7@alumnes.ub.edu

Wendy Elvira-García
Department of Spanish Language and General Linguistics
UNED, Spain
welvira@flog.uned.es

Abstract

Despite the increasing number of studies in L2 prosody, little research has been carried out on the Chinese-Spanish language pair. This article sets out to examine the L2 acquisition of nuclear contours and pitch implementation details of Spanish spoken by Chinese speakers. To this end, 555 utterances (produced by 37 informants) were analyzed within an autosegmental-metrical framework, and pitch values were evaluated using long-term distributional (LTD) and pitch dynamism quotient (PDQ) measures. The results suggest a hierarchy of difficulties in acquiring the prosodic features of different sentence types. The most salient intonational error made by the Chinese learners was the tendency to replace low nuclear accents with high/rising tones. Furthermore, the higher pitch level, narrower span, and lower F0 variance found for Chinese speakers lend support to previous hypotheses which proposed a general pitch compression pattern for L2 speech. Nevertheless, with increasing proficiency in Spanish, learners appear to develop more target-like intonation contours and pitch profiles. Finally, gender and stress effects as well as other interactions prove that L2 prosody learning is more complex than previously stated, and is influenced not only by the L1 system and oral competence but is also correlated with some psychological and sociocultural factors.

Key words: Nuclear contour; pitch characteristics; L2 speech acquisition; prosodic transfer; Peninsular Spanish by L1 Chinese speakers.

Resumen

A pesar del creciente número de estudios sobre L2, el par de lenguas chino-español ha sido abordado en escasos estudios. Este artículo examina la adquisición de los contornos entonativos y los detalles de implementación tonal del español por parte de los sinohablantes. Para hacerlo, se han analizado 555 oraciones producidas por 37 informantes usando la teoría métrico-autosegmental, y se han evaluado acústicamente sus valores tonales a través de las medidas de distribución a largo plazo (LTD) y coeciente de dinanismo de tono (PDQ). Los resultados sugieren una jerarquía de dificultad en la adquisición de los patrones entonativos relacionada con los diferentes tipos oracionales. Además, el error entonativo más destacado para los sinohablantes es la tendencia a reemplazar los acentos nucleares bajos por tonos altos/ascendentes. Además, nuestros resultados apoyan las hipótesis que proponen una tendencia general de compresión tonal en la L2, ya que se ha documentado un nivel tonal más alto, un rango más reducido y una variación menor de F0 en los sinohablantes que en el grupo control. Sin embargo, los aprendices parecen desarrollar contornos entonativos y perfiles tonales más parecidos al objetivo a medida que aumenta la competencia en L2. Finalmente, el efecto del género y acento, así como otras interacciones, demuestran que la adquisición de la prosodia es más compleja de lo expuesto anteriormente, y que se ve influenciada no solo por el sistema de la L1 y las competencias orales, sino también por factores psicológicos y socioculturales.

Palabras clave: Contorno nuclear; característica tonal; adquisición del habla de L2; tranferencia prosódica; español peninsular por los sinohablantes.

1. Introduction

Prosody in second language acquisition (SLA) research has long been promoted through a monolingual linguistic ideology. The vast majority of SLA studies on phonetics and phonology, to date, overtly or covertly tend to treat native speakers as a yardstick for second language (L2) learning. In this regard, L2 learners are defined as failed and deficient monolinguals that will not achieve a native proficiency because of their unchangeable conditions at birth (Cook, 2012; Ortega, 2010, 2011, 2014). However, since the mid-1990s, there has been continual criticism against this dominant paradigm and the concept of monolingualism as a default norm for SLA. For instance, Ortega pointed that the major deleterious consequence of the monolingual view is that it unethically turns bilinguality into an invisible reality by "erasing bilinguals' other language competence from analysis" (2010: 56). To overcome this pervasive monolingual bias, a new trend of bi/multilingual turn has been developed in recent years (May, 2013; Ortega, 2010, 2011, 2014; Valdés, Poza & Brooks, 2015). The creativity

of this new paradigm is that it conceives SLA as a dynamic and constantly interactive state between the first (L1) and the second language, rather than the sum of two clearly separate monolinguals (May, 2013). Besides, instead of examining SLA development based on native speakers, it proposes to investigate learners' total language repertoires, and utilize fully-developed bilinguals or successful L2 users as a new empirical baseline for L2 learning (Cook, 2016; Ortega, 2010, 2014).

Despite the advantage of a bilingual approach portraying a more complete picture for SLA research, it has not been sufficiently worked out to replace the predominant SLA theories, at least in the domain of phonetics and phonology. The main causes of this, according to May (2013), have been correlated with the ideological roots of monolingualism as the normative condition of the discipline. Another possible reason lies in the ambiguity and complexity of correctly defining the mature and experienced L2 users as the reference for SLA (Cook, 2016). Moreover, the lack of integration between theories and practices in different SLA contexts and fields also have stunted the epistemological change toward bilingual turn in SLA (Pandey, 2016). Finally, as mentioned by Pandey (2016), cross-continental examples or typologically distinct language pairs are still needed to explore the global appeal of current frameworks on the bilingual turn. Given these remaining uncertainties in the bilingual turn and the typological distance between Chinese and Spanish prosody (see the discussion in section 2.1), we decided in this study to investigate SLA from the classical view of native speakers. Unlike Spanish and most Romance languages that are intonation languages, Chinese is widely known as a tonal language, with each tone having a different internal pattern of rising and falling pitch contours (Yip, 2002). Arguably, problems would arise when comparing the intonation of Chinese with Spanish, given that the pitch contours of Chinese are highly dependent on the tone type of the sentence-final syllable. Therefore, rather than direct comparisons across learners' language pairs, it seems to be more feasible in our research to interpret the L1 influence on L2 acquisition based on previous empirical studies on Chinese prosody. The L2 transfer to L1 prosody from the opposite direction will be considered for further research.

Cross-language transfer effects have been discussed over the last decades in SLA research and most studies have focused on segmental differences across languages. However, several intonational aspects, for instance, F0 register, pitch range, and intonational pitch patterns are, in essence, more vulnerable to cross-language influences, and therefore more difficult to learn, to interpret, and to investigate in SLA (Mackay, 2000; Atoye, 2005; Mennen, 2014, 2015). These difficulties in L2 prosody are not merely caused by the typological distance between the phonological and phonetic systems of the L1 and the L2, but are also closely linked to the "complexity and multidimensionality" of intonation (Mennen & Leeuw, 2014: 187). Intonation

can interact with other prosodic structural elements (e.g. duration, rhythm, intensity and lexical stress) and signal multiple pragmatic functions in the speech. Therefore, it is difficult to determine whether some intonational variations in the phonetic implementation are either categorical or gradient to the phonological representations (Mennen, 2015; Nolan, 2006). Besides, the divergence of different types of frameworks on prosodic typology (e.g. see the discussion between Beckman & Venditti, 2011; Hyman, 2006, 2012; Jun, 2006; Ladd, 2001) adds to the existing difficulty of describing and characterizing intonational features. This methodological issue remained unsolved in the prosodic field until the advent of the AM (autosegmental-metrical) theory that has allowed researchers to analyze and uniformly compare the intonation systems of many languages.

The broadly recognized autosegmental-metrical framework was developed around the core idea of isolating the categorial phonological elements from its surface phonetic realization (Gussenhoven, 2004; Ladd, 1996; Pierrehumbert, 1980). This separation between the phonological and phonetic dimensions is of great importance, not only for the prosodic investigation of a wide variety of languages (e.g. Hualde, 2003; Face & Prieto, 2007; Estebas-Vilaplana & Prieto, 2008, 2010 for Spanish; Tseng, Huang, & Beckman, 2011 for Mandarin Chinese), but also for cross-language or cross-dialectal comparisons of intonation (e.g. English and Spanish comparison by Bowen, 1956 and Vilaplana, 2008; Mandarin and English by Crosswhite & McDonough, 2000; Spanish and Catalan by Vilaplana, 2008; Majorcan and Minorcan Catalan by Payà & Vanrell, 2005). According to the AM theory, intonational pitch contours are phonologically comprised of two types of tonal units: pitch accents and boundary tones, which are realized as either a high or low tone attached to metrically prominent syllables or the edges of prosodic phrases. Different languages may differ typologically in the inventory of pitch accents, boundary tones and combinations, using strategies of the structural elements according to the text structure (Graham & Post, 2018; Ladd, 1996; Mennen, 2015). Beyond this, the surface phonetic shapes (e.g. pitch height, pitch span, alignment, speech rhythm, etc.) of the intonation primitives (basic components) may also vary to different degrees because of different conventions of pitch implementation across language communities.

Furthermore, in the recently developed L2 Intonation Learning Theory (LILt), Mennen (2015) suggested that there can be some cross-language intonation differences in the semantic and frequency dimensions, not merely in the systematic (or 'phonological') and realizational (or 'phonetic') aspects. Concerning deviations in the semantic dimension, Chinese learners of L2 English were found to have difficulties in signaling new information (Juffs, 1990), marking contrastive stress (Wennerstrom, 1998) and prominence relationships (McGory, 1997) in a native-like way, even

having reached a high level of proficiency in the L2. Besides, significant contrasts were observed in the frequency of usage of phonological tones in the L1 and the L2, probably due to the transfer of the first language (see examples in Backman, 1979; Hewings, 1995; Jilka, 2000; Mennen et al., 2010). Based on this multi-dimension modeling, L2 intonation research is expected not only to examine learners' phonetic and phonological deviations from the target language but also to shed new light on their ability to use appropriately different categorical types and their ability to use intonation to signal different functional meanings in the cross-language speech.

Overall, the current study was undertaken based on the deep phonological understanding of intonation (AM) and previous evidence developed in the L2 prosody learning model (LILt). It was of interest to investigate the systematic acquisition of the intonation contours and pitch implementation details in five question types with different linguistic meanings in Spanish by L1 speakers of Mandarin Chinese and Peninsular Spanish. Furthermore, this study examined the correlation between different pitch variables and compares prosodic performance across different proficiency levels, question types, stress positions, and gender. The rest of the paper is structured as follows: Section 2 describes the intonation system of Peninsular Spanish and Mandarin Chinese and reviews the existing literature on pitch range implementation. Beyond this, the cross-language differences between L1 and L2 prosody will be introduced in this section. Next, Sections 3 and 4 describe the present study with the specific research questions and the methodology. Sections 5 and 6 report the experimental results and discuss the main findings of the work. Finally, in Section 7, we present the conclusions, the pedagogical implications, and some potential limitations of the current study.

2. Literature review

2.1. Intonation contours in L1 and L2

As Spanish is an intonation language, there is a general consensus on the use of ToBI (Tones and Break Indices) labeling systems for the description of intonational phonology. ToBI conventions distinguish two types of tonal events (pitch accents and boundary tones) and describe the intonational patterns by means of two tones that indicate prosodic level (e.g. low is L, high is H). If the tone is associated with a prominent or metrically strong syllable, the letter is followed by a star (*), for example, L* stands for a low stressed syllable. In Spanish, if the prominent syllable has more than one tone attached, the tones are linked with a plus sign and the star indicates the position of the stress. For example, L+H* signals a rising stressed syllable, whereas

L*+H signals a low stressed syllable and a high poststressed syllable. The group of tones associated to a strong syllable is called a pitch accent. If the tone is associated with the edge of an intonation group, it is marked with a percentage sign (%) when the intonation group is final (signaling an intonation phrase), or with a dash (-) when it is non-final (signaling an intermediate phrase).

Based on the ToBI framework, 2 monotonal pitch accents (L*, H*), 3 bitonal pitch accents (L+H*, L+¡H*, H+L*), 2 monotonal boundary tones (L-/L%, H-/H%) and 2 bitonal boundary tones (LH%, HL%) (see the schematic representations in Figure 1) have been proposed for the nuclear configuration of questions in Peninsular Spanish (Estebas-Vilaplana & Prieto, 2010; Estebas-Villaplana, 2008; Face & Prieto, 2007; Hualde & Prieto, 2015; Prieto & Roseano, 2019). While the monotonal accent L* is commonly found in the nuclear position of interrogative modality (e.g. in the nuclear location of information-seeking *yes-no* questions, information-seeking *wh*-questions, confirmation-seeking tag questions as well as the final constituent of disjunctive questions), the high pitch accent H* often appears in the nuclear position of echo and rhetorical questions (Estebas-Vilaplana & Prieto, 2010). With regard to the bitonal accents L+H* and L+¡H*, these are frequently linked to obviousness meanings or some nuance of surprise and exclamation, and normally appear in statements of the obvious, exclamative statements and various counterexpectational questions of Peninsular Spanish. The phonetic distinction between the two accent types typically lies in the F0 scaling, with L+¡H* exhibiting a step rise into the highest peak of the utterance. In contrast with the two previous rising accents, in Peninsular Spanish a falling pattern (e.g. the H+L* L% found for confirmation-seeking *yes-no* question) is frequently used to express "the degree of commitment that a speaker has in the truth-value of a proposition" (which technically has been called the epistemicity) (Roseano, González, Borràs-Comes, & Prieto, 2016: 3). Nevertheless, the epistemicity encoded by the intonation strategy is generally considered to be weaker than that marked with confirmative tags, for example, ¿*no*? (no?), ¿*verdad*? (is it?), ¿*no es cierto*? (isn´t it?). In Spanish, these syntactic structures are typically pronounced with a low nuclear accent L* followed by a high-rising boundary tone H%. Aside from tag questions, the H% or H- edge tone is also attested in information-seeking *yes-no* questions and the non-final constituents of disjunctive questions. Conversely, the low-falling boundary tone L% is usually observed at the end of information-seeking *wh*-questions and disjunctive questions as well as confirmation-seeking *yes-no* questions.

Figure 1. Schematic representations of the pitch accents and boundary tones of Peninsular Spanish questions presented in our study (adapted from Estebas-Vilaplana & Prieto, 2010: 19-21)

Nuclear pitch accents

| L* | H* | L+H* | L+¡H* | H+L* |

Boundary tones

| L-/L% | H-/H% | HL% |
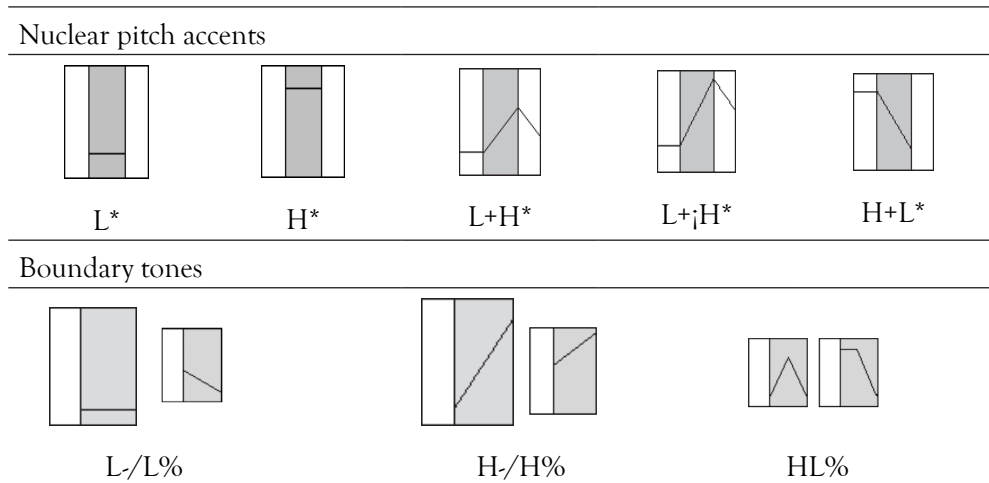
In Mandarin Chinese, however, the realization of intonation types differs strongly from intonation languages such as English and Spanish, where pitch contours are used to convey linguistic meanings and communicative purposes (Prieto & Roseano, 2019) although it is worth noting that pitch variation is typically accompanied by changes in other prosodic features, like duration, amplitude, or voice quality. As a tone language, on the one hand, the key issue in Chinese prosody research is concerned with the interaction between tone and intonation. It has long been established that Mandarin Chinese has four lexical tones and one neutral tone in its inventory, namely, tone1 (high), tone2 (rising), tone3 (falling-rising), tone4 (falling), and tone0 (phonological neutral). Though both tone and intonation use F0 as a primary cue, the functional load of F0 in Chinese lies mostly in word distinction (Chen & Gussenhoven, 2008; Yuan & Shih, 2004). Thus, it is expected that intonational meanings in Mandarin Chinese might be recognized through the modification of other prosodic indicators (normally, pitch range, pitch level, and duration), due to the restriction in manipulating the F0 curve (Chen & Gussenhoven, 2008). On the other hand, despite the growing body of knowledge on tone-intonation interaction (Chao, 1968; Chen & Gussenhoven, 2008; Girding, Zhang, & Svantesson, 1983; Kochanski & Shih, 2003; Wu, 1982; Xu, 2005, 2015; Xu & Wang, 2001; Yuan, 2004, 2006; Yuan & Shih, 2004), there seems to be no consensus as to the formulation of a general framework for Chinese prosody. Various models have been proposed to investigate the Chinese intonation mechanism, the most influential ones being the following: the PENTA model (Xu, 2005; Xu & Wang, 2001), the STEM-ML model (Kochanski & Shih, 2003) and the

StaticTarget model (Shih, 1988). Although models differ in the way they interpret and analyze intonation and its interaction with lexical tones, it is generally accepted that the intonation of Chinese questions and statements diverges increasingly at the sentence-final location. These observations are generally in line with the ToBI-based intonation school who tends to differentiate intonation types through the final nuclear configuration. Nevertheless, the high (H-/H%) and low (L-/L%) boundary labels used in ToBI systems were found to be problematic for the prosodic annotation of Mandarin Chinese (e.g. Pan-ToBI or C-ToBI system), particularly in the intonation of questions, due to the fact that the surface F0 contour of Chinese questions could have either a rising or falling end depending on the tonal identity of the last syllable

The typological differences of prosodic structure and intonation mechanism between the tone languages and the intonation languages could pose significant challenges for L2 learners of tone languages over the course of learning an intonation language. In the last decades, with the growing body of literature on L2 prosody acquisition, various intonation errors have been reported for Chinese learners of L2 English. For instance, it has been found that Chinese speakers of English tend to employ systematically high-level tones (H*) in target nuclear accents where the phonetic realization is characterized by a low plateau (henceforth L*) during the stressed syllable (Barto, 2015; Hong, 2012; Ji, 2010; Ji et al., 2009, 2012; Shao et al., 2011; Xu, 2009). The tendency to use high-level tones in the L2 nuclear position has been mostly correlated with the lack of a steady low tone (L*) in the tonal inventory of Chinese (Ji et al., 2009; Yuan et al., 2019). Additionally, the intonation stress in Chinese is achieved by mainly raising the H target tone rather than changing the low tone. Therefore, as suggested by Ling (2003) and Jia, Xiong & Li (2005), Chinese learners may unconsciously transfer the high tone strategy from their L1 to the L2 target intonation patterns when speaking an intonation language like English. Compared to the numerous findings regarding L2 English intonation produced by Chinese speakers, little research has been carried out in the case of language contact between Spanish and Chinese. To our knowledge, the only empirical investigations carried out have been those by Cortés Moreno (1997, 1999, 2001, 2004, 2005) and Liu (2005) on the L2 acquisition of Spanish prosody by Mandarin-speaking Taiwanese students. Specifically, Cortés Moreno (1997, 2004) found that Taiwanese students had considerably more problems producing the L2 Spanish intonation than perceiving it, even advanced learners who had shown native-like performance in perceiving the target intonation patterns. Moreover, studies by Cortés Moreno (2001, 2005) seem to suggest a hierarchy of difficulties in the acquisition of L2 Spanish intonation types whereby *yes-no* questions were the most difficult pattern to learn in comparison to statements and *wh*-questions. This ranking of learning difficulties holds essential implications for L2 prosody teaching. However, it is interesting to note that the said

proposal was the opposite of the sequence put forth by Yuan et al. (2019). In this latest study, statements were reported to be the most difficult L2 patterns for Chinese students due to their less perceptually salient target pitch movements compared to *yes-no* questions and requests, which involve more H-L contrasts.

2.2. Cross-language research of pitch range variation

Aside from the phonological system, languages may also differ typologically in the way they use pitch to phonetically implement the categorical elements. According to Ladd (1996), Cruttenden (1997) and Gussenhoven (2004), there are two main types of pitch range variation, namely, level and span. Whereas pitch level, or "register" has been referred to as the overall height of a speaker's voice (Cruttenden, 1997), span, or "F0 range" has been described as a speaker's "excursion size" of range of frequencies in the speech ('t Hart, Collier & Cohen, 1990: 75). A positive correlation exists between the two dimensions of pitch range; the higher the F0 level, the wider the pitch span (Urbani, 2012). In Gussenhoven's view (2004), the phonetic implementation of pitch values has generalizability in the paralinguistic intonational meanings and can be explained using a number of biological codes. For instance, the Frequency Code, which is based on the correlation between the size of the larynx and the rate of vocal fold vibration, suggested that a higher pitch frequently signals a smaller larynx and expresses speaker uncertainty, whilst a lower F0 tends to be associated with a larger organ of production and an assertive interpretation (Ohala, 1984, 1994) .

Based on the universality of biological codes, however, some languages were found to go against the general form-function patterns and show some language-specificities in the use of pitch codes. Evidence for this tendency of cross-language differences can be found in a growing number of recent studies in which different languages or language varieties have been reported to have different levels of pitch range and F0 variability in the speech. A crucial difference was found, for example, between tone languages such as Mandarin Chinese and stress languages like English. Compared to English, Mandarin Chinese shows generally higher pitch level, greater F0 fluctuation, and faster pitch change rate in the prose passage (Keating & Kuo, 2012; Eady, 1982) as well as wider F0 span in broadcast news speech (Yuan & Liberman, 2014). Similar F0 features of higher register and wider span were also reported in the Chinese dialect Min when compared with American English (Chen, 2005). Nevertheless, in Cantonese-English bilingual children's speech, significantly lower values of speaking fundamental frequency and F0 range were found for Cantonese speakers than for English-speaking children (Ng, Hsueh, Leung, & Shing, 2010). This difference in pitch implementation might be correlated with the tonal structure of Cantonese, although language proficiency and sociocultural conventions may also play a role in the

divergence of voice pitch characteristics. Furthermore, a recent study on cross-language comparison suggested a wider F0 range and higher F0 register in native Chinese speech than in native Japanese speech (Shi, Zhang, & Xie, 2014). However, when Japanese was compared to American English, Spanish and Dutch, it was found to have the highest values of mean F0 in read speech (Hanley, Snidecor, & Ringel, 1966; Van Bezooijen, 1995). The higher pitch values found for Japanese speakers, particularly for Japanese women, are not due to mere physiological or anatomical differences across gender and language communities (Rendall, Vokey, & Nemeth, 2007; Van Dommelen & Moxness, 1995), but are more closely linked to the social-cultural behaviors and relative powerlessness image of Japanese women (Van Bezooijen, 1995). Aside from the above-mentioned cases, further support for cross-language pitch value differences can be found in a large number of studies on distinct language pairs (see Mennen et al., 2014 and Ordin & Mennen., 2017 for a short review).

Compared to the accomplishments of cross-language research, few studies have shed light on L2 speech deviations from the target language. Because the acquisition of L2 intonation always entails some degree of interaction between the L1 and L2 systems, it is expected that the cross-language differences in the F0 register and span may also appear in the L2 production. Generally, it is suggested that L2 learners have a compressed F0 span and less variable pitch when compared to native speakers of the target language. For instance, Chen (1972) and Juffs (1990) reported that Chinese L2 learners had a narrower F0 span than native speakers of English. Recently, using a large-scale dataset extracted from a language learning app, Yuan et al. (2018) reconfirmed that, compared to that of native English speakers, the speech of Chinese L2 learners was characterized by a narrower pitch span, slower pitch change rate and more small "ripples" on the F0 contour. In addition, it was found that Chinese learners of L2 English and L2 German had higher values than native speakers in F0 span on the phoneme level, and in pitch change amount on the utterance level, due to the negative influence of L1 mandarin prosody (Ding, Hoffmann, & Hirst, 2016; Ding, Jokisch, & Hoffmann, 2012). This general trend of compressed range and less F0 variability in L2 speech can be observed in many L1-L2 combinations, for example, in Spanish learners of L2 English (Backman, 1979), in Chinese learners of L2 Japanese (Shi et al., 2014), among many others. The consistency seen in L2 pitch implementation patterns is probably influenced by the L1 prosody, but more frequently it has been correlated with learners' lack of confidence or cautiousness when speaking a non-native language (Mennen, 1998; Shi et al., 2014; Volín, Poesová & Weingartová, 2015). Another plausible reason for those L2 speech deviations could be that learners are too focused on the segmental pronunciation and stress emphasis, thus, there might be a lack of attention given to extending and varying the F0 pitch in a native-like way (Zimmerer et al., 2014).

In light of all these findings, it seems that there is a universal trend of pitch range compression when speaking an L2. However, in previous studies, most experiments were conducted with a small number of subjects. As a result, the conclusions drawn are potentially less convincing due to the high variability of the F0 range within speakers (Ladd, 1996). Another problem concerning previous studies is that in some cases the pitch range has been treated as a unitary concept without distinguishing the level from the span (Mennen et al., 2014), and have analyzed the data using different quantification measures. Most importantly of all, to date, none of the studies have explored the L2 pitch implementation characteristics of Spanish by L1 Chinese learners, thus, this research would be interesting evidence for the investigation of a general pitch compression pattern in the L2.

3. The present study

As can be seen in the literature review, prior studies in L2 prosody learning give inconsistent evidence for the cross-language differences in the phonological and phonetic dimensions, either due to the different speaking materials and quantification methods used in the research or because of the different language pairs under investigation. Of these studies, few have shed light on the production of Spanish intonation by L1 Chinese speakers, probably owing to the typologically substantial differences between the two language systems. Therefore, with the present study we intend to fill the existing gap in second language research and investigate the acquisition of intonation contours and pitch implementation details in L2 Spanish, by taking into account proficiency level (or L1), question type, gender and stress position. Specifically, the current study addresses the following questions:

1. Is the L1 prosodic system transferred (either positively or negatively) to the L2 Spanish intonation and therefore can the L1 account for some L2 deviation errors? If so, does the acquisition of L2 intonation patterns reflect different levels of proficiency?

2. Does the acquisition of pitch implementation details (as measured by six pitch variables based on the F0 distribution: mean F0, max F0, min F0, 100% span, 80% span and Pitch dynamism quotient -PDQ-) in an L2 reflect different levels of proficiency, and do they differ among different question types or stress positions as well as between male and female speakers?

3. Do our findings of L2 pitch range variation point towards a universal developmental trajectory (narrower span and less variable pitch) during the

L2 learning process, or are they highly dependent on the L1-L2 language pairs under study?

4. Does the acquisition of L2 Spanish intonation show different levels of difficulty depending on pragmatically different question types, and, if so, does this difficulty ranking exist only in the phonological dimension, or it can also appear in the phonetic dimension?

4. Methodology

4.1. Participants

The subjects of the present study were 5 native speakers of Peninsular Spanish and 32 learners of Spanish with Mandarin Chinese as their first language. The age of the participants ranged from 18 to 31 years (mean age: 23.97; SD=2.872). None of the individuals reported any speech, hearing or communicative impairments. The native control group consisted of 5 women who were born and/or lived for more than 20 years in Barcelona, and who had a comparable level of education (mean age: 23.2; SD=4.87). Although some of these participants were Catalan-Spanish bilinguals, they reported that Spanish was their dominant language.

As for the Chinese speakers (26 females and 6 males), they were all students and lived in Barcelona at the time of the recordings (mean age: 24.09; SD=2.53). Peninsular Spanish was the language variety (dialect) to which they had been predominantly exposed both during their leaning period in China and their immersion period in Spain. The L2 participants were divided into 2 groups according to proficiency level in line with the Common European Framework of Reference for Languages (CEFR): intermediate level (B1-B2) and advanced level (C1-C2). The Spanish language proficiency of the Chinese speakers was judged using the official language qualification DELE (Diploma of Spanish as a Foreign Language), with the exception of those learners who did not have this certificate. In the latter case (roughly 15% of the L2 learners), participants were required to state their self-evaluated L2 proficiency on the basis of Spanish language courses they had completed. In order to ensure that the learners were aware of the criteria of self-assessment, explicit descriptions of the six levels of European language proficiency were explained to those speakers at an early stage of this process.

In this study, we did not specifically control for Chinese learners' origin, age of L2 acquisition, or length of exposure to the target-language environment, due

to the dramatic reduction in subject pool which would result from including these selection criteria. However, as these variables were reported in previous literature to exert certain effects in the L2 speech (Cadierno et al., 2020; Juan-Garau & Pérez-Vidal, 2007; Kharkhurin, 2008; Pfenninger & Singleton, 2016), we decided to include this additional information for the non-native groups (see Appendix A for more details). The L2 participants were native speakers of Mandarin Chinese and declared that this was their dominant language, despite having different places of origin within China. The majority of the learners acquired Spanish in adulthood (mean age: 18.81; SD=2.08), only two female participants reported being teenage learners (they started to learn Spanish at 12 and 17 years old). Although the number of months of exposure to the target language was quite different between individual learners, the mean exposure time of advanced speakers (mean length:19.13; SD=9.51) was generally longer in comparison to the intermediate group (mean length: 22.8; SD=18.02).

4.2. *Materials and recording procedures*

In order to collect natural speech, we used the DCT (Discourse Completion Task) (Billmyer & Varghese, 2000; Félix-Brasdefer, 2010; Golato, 2003) method to elicit the corpus. Specifically, 15 daily scenarios were designed to elicit five question types with different linguistic meanings in the target language, namely, information-seeking *yes-no* questions ('YN'), information-seeking *wh*-questions ('WH'), disjunctive questions ('DJ'), confirmation-seeking *yes-no* questions ('CYN') and confirmation-seeking tag questions ('TAG'). Each question type varied in nuclear stress position (two positions: final and penultimate stressed syllable). Test items were mostly comprised of words with high familiarity ratings or high frequency (Tanaka & Terada, 2011), for the benefit of non-natives' comprehension during the task activity (see Appendix B for more details). The average syllable number in the study was 5.8 per utterance.

Situational contexts were presented by an interlocutor with whom the participants were somewhat familiar, and speakers were asked to produce the target sentence used in that situation. The task was performed only once except in cases where there was a problem with the speaker's first realization. Subjects were allowed to reproduce the test item if they made a mistake. All recordings took place in a soundproof room with a head-mounted microphone. Speech files were digitized at a sampling rate of 44.1 kHz and with a quantization precision of 16 bits. Each utterance was saved separately as a *wav* format file and annotated to a *TextGrid* object using a *Praat* script.

4.3. Data collection and analysis

4.3.1. Intonation contours

For the benefit of intonation labeling, unvoiced segments were interpolated through and F0 trace was smoothed using a *Praat* script with bandwidth set to 10 Hz. In our study, intonation transcriptions were realized by combining the visual representation of the F0 curve with the auditory perception of pitch accents. To facilitate the transcription work, all test items were initially annotated using a prosodic tool (Eti-ToBI) which automatically labels intonational events in Spanish utterances (Elvira-García et al., 2016). However, as this script was developed based on Sp_ToBI and Cat_ToBI conventions for native speakers, many unexpected pitch movements produced by Chinese learners could not be appropriately assigned (mainly due to differences in pitch alignment). Thus, manual correction of all labeling was conducted by the first author of the paper, according to the guidelines for Castilian Spanish intonation (Vilaplana, 2008; Vilaplana & Prieto, 2010). Furthermore, the annotation results were checked and revised by the second author of the paper and a third expert in Sp_ToBI labeling.

To conduct a cross-language comparison of intonation contours, the proportion of occurrence of pitch accents, boundary tones and nuclear configurations was calculated separately within the 5 question types. In this study, we were particularly interested in the nuclear configuration realization which has been referred to as the most salient part of an intonation contour (Prieto & Roseano, 2019) although it is worth noting that pitch variation is typically accompanied by changes in other prosodic features, like duration, amplitude, or voice quality. Regarding the set of linguistic functions that intonation (together with other prosodic factures, other parts of the F0 contour were excluded from further analysis.

4.3.2. F0 *pitch extraction*

Pitch tracking was performed automatically in *Praat* using the ESPS algorithm ('get F0') (Talkin, 1995), with pitch floor set to 70 Hz and pitch ceiling 600 Hz. A time step of 10 ms was used for the computation of F0. The original F0 data was refined by a manual correction based on *Pitch* objects. Specifically, cases with octave jumps and measurement errors (e.g. mistakes caused by creaky voice or laryngealizaiton, or false voicing in silent fragments) were marked as "unvoiced" and therefore excluded from further analysis. The speakers' pitch characteristics were analyzed along three dimensions: (a) *pitch level*, which has been referred to as a "reference line" (or a constant value) averaged by the rising and falling movements in the overall pitch contour

(Gussenhoven, 2004: 76); (b) *pitch span*, that is the distance between the highest and lowest values of the F0 contour, and finally, (c) *pitch variation*, which describes the degree of F0 variability in the speech (J't Hart, Collier & Cohen, 1990).

To quantify pitch characteristics in Spanish L1 and L2, two types of measures were included in the analysis. Firstly, for pitch level and span, we used the long-term distributional (LTD) measures developed by Mennen (2014). Specifically, 3 parameters have been extracted for level, and 2 parameters for span including both the absolute pitch excursion (100% span: max F0 - min F0) and the 80% span (the 90th and 10th percentile span) which has been reported to be the best F0 distributional measure (Mennen, Schaeffler, & Docherty, 2009; Niebuhr & Skarnitzl, 2019). Furthermore, because female and male speakers differed greatly in group size, we decided to introduce another pitch dynamism quotient measure (PDQ) to normalize the F0 variation data and to make the LTD results more robust. PDQ values were calculated by dividing the standard deviation by the mean frequency of each utterance. This metric gives an account of pitch variability and generally, the higher the PDQ, the greater the F0 variance and the emotional turmoil in subjects' speech (Shi, Zhang, & Xie, 2014; Wang & Qian, 2018; Zimmerer et al., 2014). To sum up, the following six dependent variables were extracted for the comparison of L1-L2 pitch implementation characteristics:

- Pitch level: *min F0*, *max F0* and *mean F0*
- Pitch span: *100% span*, *80% span*
- Pitch variability: *PDQ*

To assess the effect of scales, pitch measures of the span were also transformed into another two psycho-acoustic scales. Compared to the corresponding linear scale (Hz), the logarithmic (semitones) or the near-logarithmic (ERB-rate) scales have been reported to be the best measures for modeling intonational equivalence between females and males, and for capturing the frequency differences across speakers and languages (Nolan, 2003; Patterson & Ladd, 1999). Besides, correlation coefficients were calculated using the Pearson method in order to examine whether, and to what extent, F0 span is correlated with other pitch variables.

4.4. Statistical analysis

During the first stage, the between-group differences with regard to the production of pitch accents and boundary tones were tested in the R environment (R Core Team, 2016) using the chi-square statistic with continuity correction when all the cells of

the contingency tables for expected values had a value greater than 5. In the case that this condition was not met, we applied Fisher's exact probability test to examine the significance. At the second stage, statistical analysis of pitch results was performed using the linear mixed effects model. The model was run separately for six dependent pitch variables (namely, min F0, max F0, mean F0, 100% span, 80% span and PDQ) with Proficiency, Gender, Question type and Stress position, as well as all their possible interactions as fixed effects. Subjects were included as random effects with all possible random intercepts. We used the Anova function to test the significance of main effects, and *p* values were fitted by eliminating non-significant effects of the initial model with the *lmerTest* package and evaluated using Satterthwaite's approximation (Kuznetsova, Brockhoff, & Christensen, 2017). The post-hoc comparisons were conducted using the single-step function of the *multcomp* package (Hothorn et al., 2016) supported by the *emmeans* function (Lenth, Singmann, & Love, 2019). Finally, the correlation coefficients of various F0 distributional measures were obtained using Pearson's method (Benesty et al., 2009).

5. Results

5.1. Results of overall pitch accents and boundary tones

The first analysis examines the proportion of pitch accents and boundary tones of the five question types within each of the three language groups, namely, L2 Chinese intermediate learners (hereafter CI), L2 Chinese advanced learners (hereafter CA), and L1 Native Spanish speakers (hereafter SN). Of the 555 speech items, 119 items were realized with a low-rising accent L+H* and 61 items, with a high-level tone H* in the nuclear location, regardless of final pitch movements. Interestingly, both accent types were produced by Chinese L2 speakers, with a relatively higher proportion for the CI group than for the CA group (see Table 1). However, the probability test revealed no statistical significance between the two learner groups with regard to the production of the two pitch accents H* (χ^2=1.20, p>0.1) and L+H* (χ^2=0.72, p>0.1). The vast number of high plateaux and rising patterns in L2 Spanish corroborate previous findings for L2 English and L2 German spoken by Chinese learners (Ding et al., 2012; Ji et al., 2009, 2012). Nevertheless, concerning the nuclear accent L*, the SN group used a significantly higher number of steady low tones compared to the CI (χ^2=8.94, p<0.01) and CA learners (χ^2=5.12, p<0.05). Similar results have also been observed in the realization of the pitch accent H+L*, whereby Spanish L1 speakers presented a significantly higher proportion of falling contours than the two learner groups (CI-SN: p<0.001; CA-SN: p<0.01).

Table 1 Proportion of pitch accents produced by the three language groups over the five question types

	H*	L+H*	L+¡H*	L*	H+L*	Total
CI	14.58%	26.67%	0%	58.75%	0%	240
CA	10.83%	22.92%	0%	63.75%	2.50%	240
SN	0%	0%	10.67%	78.67%	10.67%	75

In regard to the boundary tones, Table 2 shows that, compared to the Spanish native group, Chinese L2 learners, particularly the CI group, produced significantly more rising boundaries (H%: χ^2=4.39, p<0.05) and less falling boundaries (L%: χ^2=5.99, p<0.05) over the five question types. In addition, a small proportion of rising-falling boundaries (HL%) was exclusively observed in the L2 intonation (see Table 2). These results appear to be contradictory, to some extent, to our initial expectation that L2 learners tend to use a final fall more frequently than a final rise for Spanish questions, given that in their L1 Chinese questions are marked mostly by interrogative particles. A possible explanation for this phenomenon may be that Chinese learners have overproduced the typical final rise for all question types, due to a lack of knowledge about the intonational phonology of the target language (both its possible patterns and their communicative functions). Furthermore, our study seems to suggest a general trend of improvement during the L2 acquisition of phonological tones. More precisely, it was found that compared to Chinese intermediate learners, the advanced group showed systematically a more native-like performance in producing the target pitch accents and boundary tones. However, these between-group differences are not statistically significant (p>0.1 in all cases except for the nuclear accent H+L* where the p value was less than 0.05). This is likely because the oral proficiency levels of Chinese intermediate (B2 level) and advanced learners (mostly C1 level) were very similar in our research.

Table 2. Proportion of boundary tones produced by the three language groups over the question types

	H%	L%	HL%	Total
CI	67.50%	30.42%	2.08%	240
CA	64.17%	34.17%	1.67%	240
SN	53.33%	46.67%	0%	75

5.2. Results of nuclear configurations

5.2.1. Information-seeking yes-no questions

Since pitch contours did not differ significantly between the two Chinese groups (see the statistical results in section 5.1), we treated the non-native subjects as a single group when comparing them to L1 native speakers in the nuclear configuration analysis. Specifically, for information-seeking *yes-no* questions, only around 40% of the Chinese learners successfully acquired the native-like pattern (L* H%), which was significantly less than the Spanish L1 group who consistently used the typical low-rising contours for *yes-no* questions (χ^2=16.64, *p*<0.001). The majority of L2 speakers were found to apply an early rising accent (L+H*: 35%) or use high-level accents (H*: 18%) with a rising end (H%) on the nuclear position (see Figure 2). The large number of deviated accents produced by Chinese learners may be explained as a negative transfer from their L1. Mandarin Chinese does not have a steady low tone (L*) in its tonal inventory and intonational stress is realized mainly through the pitch range expansion of high tone, therefore, Chinese learners may unconsciously transfer the high-level tone (namely Tone1) or the rising tone (namely Tone2) from their L1 into the L2 Spanish prosody. Moreover, it was noted that most of the Chinese speakers were capable of producing the final rising boundaries (H%) of *yes-no* questions, with only 7% of the L2 speakers failing to achieve the high pitch targets in sentence-final locations.

Figure 2. Waveform, spectrogram and F0 trace of the information-seeking *yes-no* question "¿Tiene mandarinas?" '*Do you have Tangerines?*' produced by a native speaker (L* H%) (left) and a Chinese learner (L+H* H%) (right)
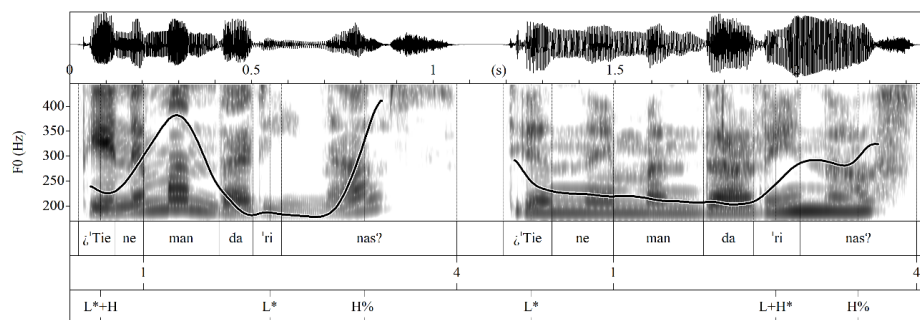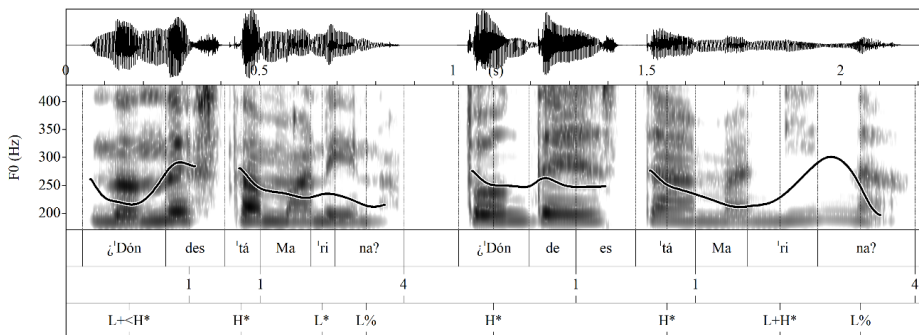
5.2.2 Information-seeking wh-questions

With regard to information-seeking *wh*-questions, the Spanish L1 speakers produced three possible pitch contours in the nuclear position: L* L% (30%), L* H% (18%) and L+¡H* L% (53%). Of these patterns, the first falling pattern (L* L%) has been interpreted as the most neutral and default realization of *wh*-questions, while the second rising contour (L* H%) expresses a nuance of politeness and the speaker's interest or greater involvement in the informative speech act (Estebas-Vilaplana & Prieto, 2010: 35). The third pattern L+¡H* L% is not an information-seeking *wh*-question, but has instead been described as a form of the reintroduction of a previously stated topic. The high proportion of this complex contour in our corpus can likely be attributed to different interpretations of the situational contexts by the L1 speakers.

As for the L2 intonation results, the Chinese learners showed a great probability of success in producing the two typical falling (L* L%: 30%) and rising contours (L* H%: 18%) used for the target *wh*-questions, and no statistically significant difference was found between the Spanish L1 and L2 groups (*p*>0.1). Aside from the native-like production, the rest of the L2 speakers were found to show deviation problems similar to those of information-seeking *yes-no* questions, that is, a tendency to apply rising tones (L+H*: 44%) or high-level (H*: 8%) tones instead of low accents (L*) in the nuclear location (see Figure 3).

Figure 3: Waveform, spectrogram and F0 trace of the information-seeking *wh*-question "¿Dónde está Marina?" '*Where is Marina?*' produced by a native speaker (L* L%) (left) and a Chinese learner (L+H* L%) (right)


5.2.3 Information-seeking disjunctive questions

Disjunctive questions in L1 Peninsular Spanish were realized mainly with a rising contour (L+H* H-: 93%) in the first prosodic unit and a falling movement (L* L%: 100%) at the end of the utterance (see Figure 4). Among Chinese learners, in the

first prosodic unit, 81% of the participants successfully produced the target rising pitch pattern (L+H* H-), statistically they showed no significant difference compared to the L1 native performance (χ^2= 0.62, p>0.1). For the second prosodic group, more than half of the L2 learners (roughly 66%) were able to produce the final falling pitch contours (L* L%) found in disjunctive questions. Nevertheless, according to Fisher's exact test, there is a significant difference (p<0.01) between the L1 and L2 speakers with regard to their intonational performance. The high level of accessibility of target intonation contours of disjunctive questions by Chinese speakers could be explained as a positive transfer of L1 intonation strategies. In Mandarin Chinese, disjunctive questions are realized in a similar way to those of Peninsular Spanish, that is, by the expansion of pitch range in the first prosodic unit and compression in the last constituent of the sentence. This similarity in the phonetic dimension between the source and target languages appears to benefit or accelerate learners' rate of acquisition of target intonation patterns. In addition to the low tone (L*) deviations, it is also interesting to note that some Chinese learners might make mistakes when producing the intermediate and final boundaries of disjunctive questions. For instance, roughly 24% of Chinese speakers were found to use a high-rising boundary (H%) instead of a low-falling boundary (L%) in the final constituent of disjunctive questions (see Figure 4). This may happen because of their lack of intonational knowledge and the cognitive bias in marking questions with final rising pitch movements.

Figure 4. Waveform, spectrogram and F0 trace of the disjunctive question "¿Quieres melón o melocotón?" '*Do you want melon or peach?*' by a native speaker (L+H* H- L* L%) (left) and a Chinese learner (L+H* H- L* H%) (right)
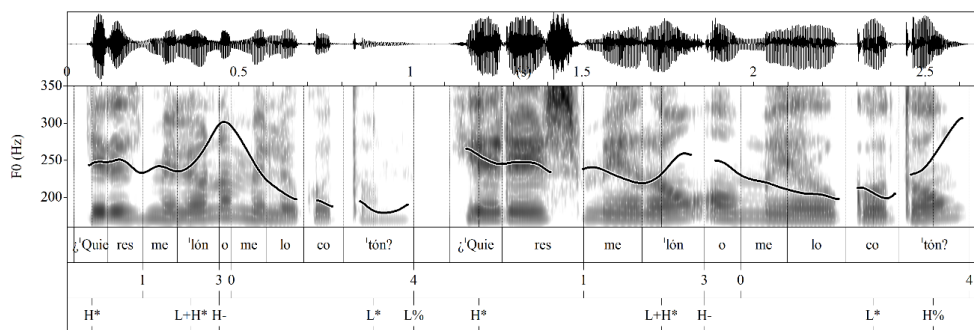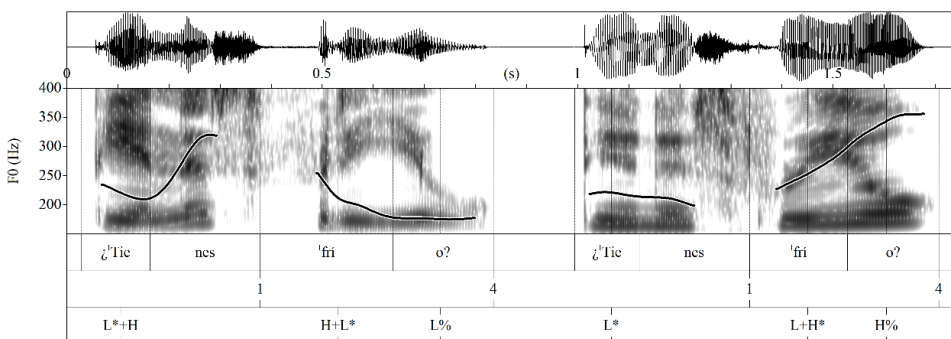

5.2.4 Confirmation-seeking yes-no questions

In addition to the prototypical informative speech act, questions in Peninsular Spanish can also be used to convey the illocutionary force of confirmation. Unlike information-seeking questions in which speakers do not have any expectations

regarding the answer, confirmation-seeking questions are biased questions where the information is previously given or shared through the context or presumptions, and the speaker usually expects a positive answer from the other interlocutor (Frota & Prieto, 2015; Vanrell et al., 2010). In our corpus, more than half of the L1 speakers applied a falling contour (H+L* L%: 53%) to express a relatively higher certainty on the proposition, whilst the rest (47%) maintained the canonical rising pitch pattern (L* H%) used for information-seeking *yes-no* questions.

As for the Chinese learners, only a small number (6%) had acquired the native-like falling pitch accent (H+L*) combined with a low boundary tone (L%) for this type of question. By contrast, 35% of the L2 speakers borrowed the nuclear contours of information-seeking *yes-no* questions, while the rest exhibited several different patterns which are congruent with a deviated form of L* H% with the low nuclear accent (L*) misproduced (see Figure 5). These results proved significantly different ($\chi^2=16.01$, $p<0.001$) from those produced by L1 native speakers.

Figure 5. Waveform, spectrogram and F0 trace of the confirmation-seeking *yes-no* question "¿Tienes frío?" '*Are you cold?*' produced by a native speaker (H+L* L%) (left) and a Chinese learner (L+H* H%) (right)


5.2.5 Confirmation-seeking tag questions

Confirmation-seeking questions can be produced, in addition to the typical falling pattern presented in the section above, with various confirmative tags (i.e. ¿no? (no?), ¿verdad? (is it?), ¿no es cierto? (isn't it?)). This syntactic resource is valid both in Mandarin Chinese and in Peninsular Spanish, as it expresses a stronger commitment by the subject to the truth value of the proposition. In our experiment, all question tags were consistently produced with a low-rising pitch contour (L* H%) by Spanish L1 and L2 speakers. A possible explanation for this phenomenon has been correlated with the (alleged) universality in marking question tags with various alternative strategies

such as final rise, final pitch range expansion, late prominence and some emphasis gestures (Cuenca, 1997; Gussenhoven & Chen, 2000). Besides, the simple structure (mostly 1 or 2 syllables) of those tags has also been proposed as a plausible factor which accounts for the high intonational accuracy in L2 tag questions. Note, however, that the production of target phonological patterns does not necessarily mean that L2 speakers have acquired a 100% native-like prosodic performance. It has been observed, for example, that the question tags produced by Chinese learners differed phonetically from those produced by the L1 native speakers in the F0 differences of the final rising movements (see Figure 6). Nevertheless, this comparison is outside the scope of this paper.

Figure 6: Waveform, spectrogram and F0 trace of the confirmation-seeking tag question "¿Sales al mercado, no?" '*You are going to the market, aren't you?*' produced by a native speaker (L* H%) (left) and a Chinese learner (L* H%) (right)
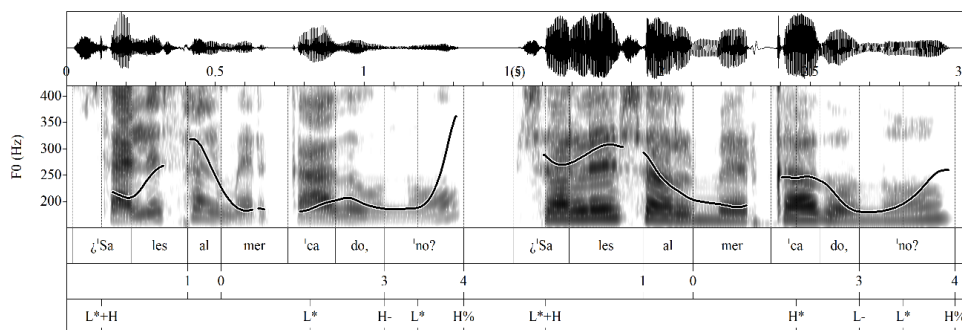
5.3. Pitch implementation results

5.3.1. Pitch level results

This section aims to show the differences in pitch level (min, max and mean F0) between Spanish L1 and L2 speakers. The statistical analysis revealed a tendency among female learners of the CI and CA groups to have systematically higher pitch levels than female speakers of the SN group (see Table 3), due to the influence of L1 tonal structure. Nevertheless, results on main effects (Table 4) show no significance of Proficiency on any of the three measures of pitch level. On the contrary, the effects of Gender, Question as well as the interaction of Proficiency and Question on the three pitch measures. Furthermore, Table 4 shows that there is a mild effect of Stress on the mean F0 and the factor Stress strongly interacted with Proficiency on the variable of mean F0. No significant interactions were found between Gender and Question, Gender and Stress on the three measures of pitch level.

Table 3. Descriptive statistics of pitch level results depending on language group and gender

|	Mean F0	Min F0	Max F0
CI_M	143 Hz	108 Hz	214 Hz
CI_F	236 Hz	189 Hz	329 Hz
CA_F	232 Hz	175 Hz	335 Hz
SN_F	220 Hz	167 Hz	327 Hz

Table 4. Effects (F values) of Proficiency, Gender, Question type, Stress position and its interactions on the three measures of pitch level (Hz) ($N=555$, '***'$p<0.001$; '**'$p<0.01$; '*'$p<0.05$; '·' $p< 0.1$)

	Mean F0	Min F0	Max F0
Proficiency	1.06	2.50·	0.17
Gender	95.01***	63.17***	37.90***
Question	8.49***	5.73***	13.38***
Stress	5.80*	0.37	3.70·
Proficiency *Question	7.80***	3.65***	6.48***
Proficiency*Stress	8.12***	0.45	0.90

On the one hand, the lack of a significant effect of Proficiency, but its strong interaction with Question type on the three F0 measures, further suggests that pitch level may differ within each of the five question types produced by learners of different levels of proficiency. Specifically, the post-hoc analysis indicates that the CI group has a significantly higher minimum (t(2)=2.59, $p<0.05$) and mean F0 values (t(2)=3.09, $p<0.01$) in CYN questions than the L1 native speakers. Also, Figure 7 shows that speakers in the CI group tend to use a higher mean pitch in WH questions than the native participants (t(2)=2.56, $p<0.05$). By contrast, the SN group was found to have a significantly higher maximum F0 in WH questions when compared to the CA group (t(2)=2.47, $p<0.05$). On the other hand, Figure 8 indicates that the two Chinese groups tend to have significantly lower mean F0 (CI: t(2)=-3.383, $p<0.01$; CA: t(2)=-4.359, $p<0.001$) in utterances with final-syllable stressed words compared to those with stress on the penultimate syllable. Moreover, the CI group was found to use a statistically lower maximum F0 in final-syllable stressed questions (t=-2.45, $p<0.05$). These results differ from those of the SN group, which exhibited higher pitch values in final-syllable stressed sentences (oxytone words), as described in Figure 8. A possible explanation for

this deviation in L2 pitch has been correlated with the extra difficulty of combining intonational patterns and those required by lexical stress, since some pitch contours require adaptations when they are applied to oxytone words (e.g. L* H% becomes L+H* L%) (Prieto & Roseano, 2019)although it is worth noting that pitch variation is typically accompanied by changes in other prosodic features, like duration, amplitude, or voice quality. Regarding the set of linguistic functions that intonation (together with other prosodic factures.

Figure 7: Mean pitch of the three language groups depending on proficiency and question type. Upper and lower levels indicate maximum and minimum F0
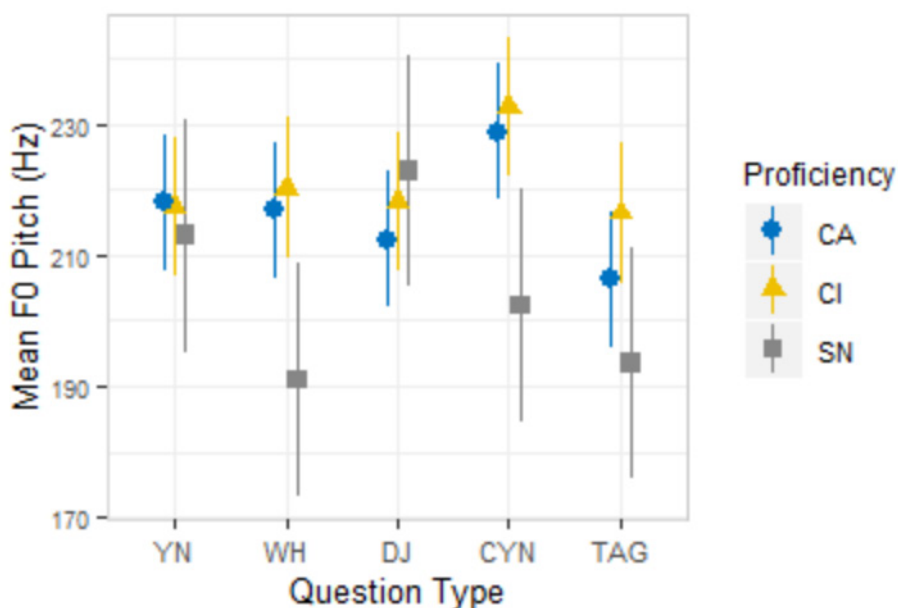
Figure 8: Mean pitch of the three language groups depending on proficiency and stress position. Error bars indicate ±1SE
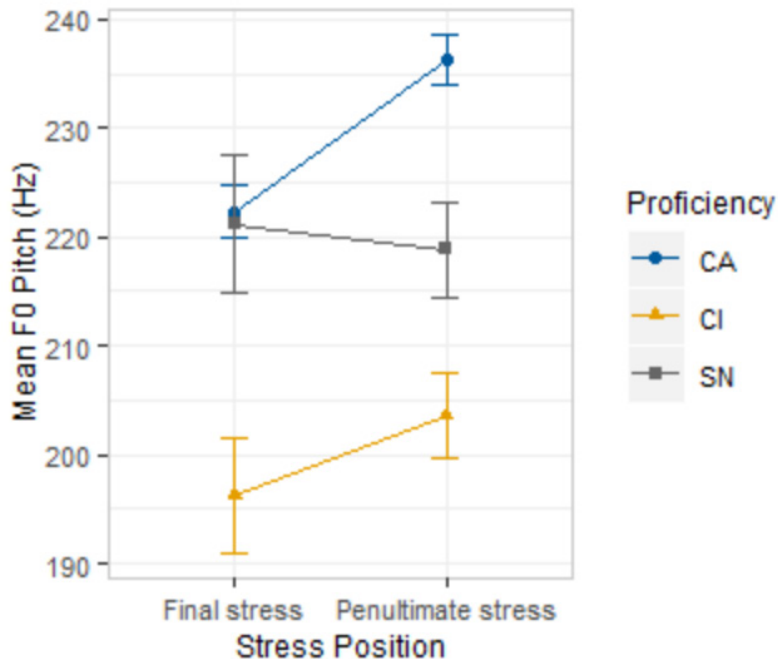

5.3.2. Pitch span results

Of the two F0 span measures, only the 80-percentile span measured in St showed a significant main effect of Proficiency ($F_{(2, 38)}=3.47$, $p<0.05$). As expected, the two learner groups showed a more reduced span than the L1 native speakers (80% span: CI:7.11 St; CA: 8.34 St; SN: 9.45 St). Nevertheless, the F0 differences were only statistically significant when comparing the CI and SN groups ($t_{(2)}=-2.53$, $p<0.04$). Besides this, there was a strong main effect of Question type (80% span: $F_{(4, 518)}=9.16$, $p<0.001$; 100% span: $F_{(4, 518)}=12.30$, $p<0.001$) as well as its interaction with Proficiency (80% span: $F_{(8, 518)}=8.78$, $p<0.001$; 100% span: $F_{(8, 518)}=7.01$, $p<0.001$) on the two span measures. Since statistical results among scales of Hz, St and ERB were similar on Question type and on the interaction between Proficiency and Question type, only the St version is reported in this section.

Furthermore, a post-hoc test which looked for the interaction between Proficiency and Question on the 80% span showed that within the five question types, the CI and CA groups tended to use a significantly narrower span than the L1 speakers in YN questions (CI-SN: $t_{(2)}=-3.38$, $p<0.01$; CA-SN: $t_{(2)}=-2.82$, $p<0.05$), DJ questions (CI-SN:

$t(2)=-3.70$, $p<0.01$; CA-SN: $t(2)=-3.53$, $p<0.01$) and CYN questions (CI-SN: $t(2)=-2.53$, $p<0.05$). By contrast, regarding WH questions, Figure 9 indicates that the two Chinese groups had a wider span than the native speakers, although this trend did not reach statistical difference in our study. Besides, the factor Proficiency was also found to interact significantly with Stress on the 80% span measured in ERB ($F(2, 518)=3.54$, $p<0.05$) and in Hz ($F(2, 518)=4.33$, $p<0.05$). More precisely, the post-hoc tests on all scales indicated that the two learner groups, particularly the CA group (e.g. St: $t(2)=-2.70$, $p<0.01$), tended to compress the span in questions with final-syllable stressed words more than those with stress on the penultimate syllable. By contrast, results in Figure 10 seem to suggest an opposite trend for the SN group regarding the span performance on the two stress positions. Finally, unlike the mean F0, no significant main effect of Gender or Stress was found on the span measures on any of the scales.

Figure 9. 80% F0 span of the three language groups depending on proficiency and question type
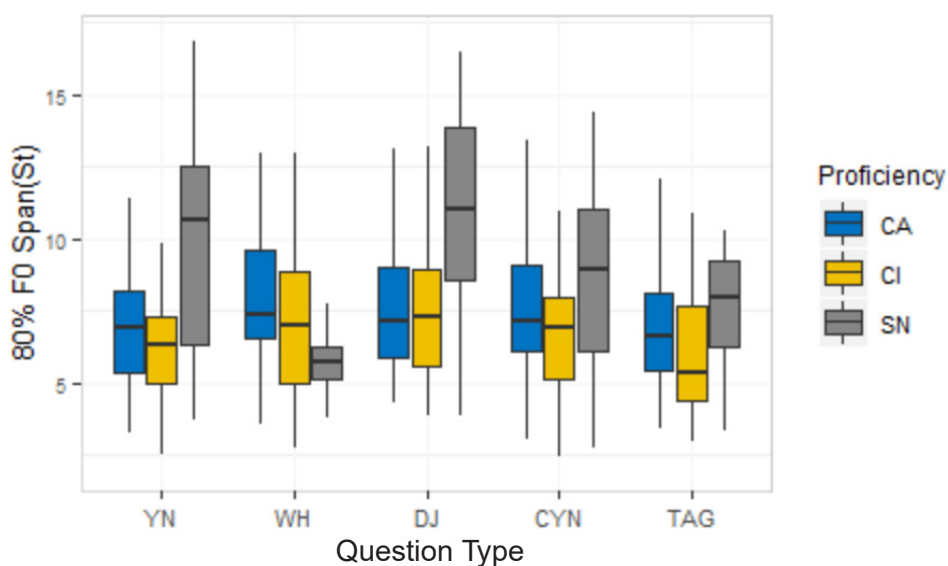
Figure 10. 80% F0 span of the three language groups depending on proficiency and stress position. Error bars indicate ±1SE
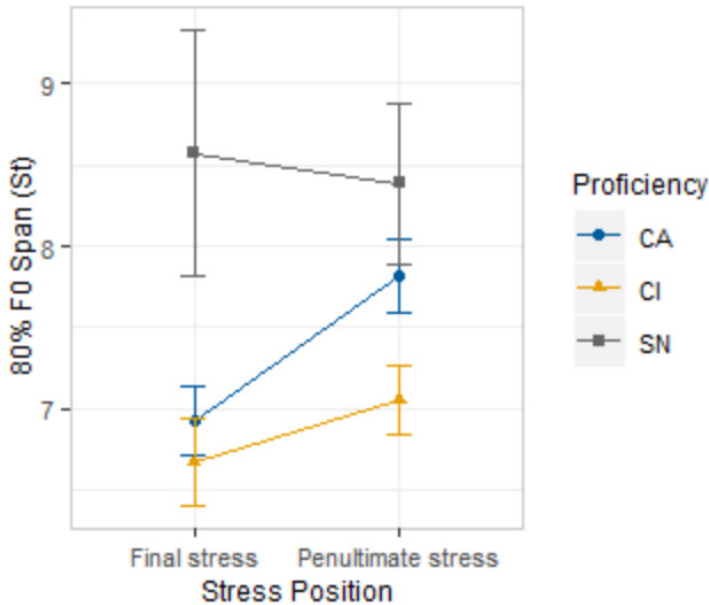

5.3.3. Pitch variation results

With regard to the pitch variation, the statistical analysis showed a significant main effect of Gender ($F_{(1, 38)}=4.56$, $p<0.05$), Question type ($F_{(4, 518)}=10.07$, $p<0.001$) and Stress position ($F_{(1, 518)}=3.98$, $p<0.05$). To be specific, our results showed that female speakers were significantly more variable than male speakers in their use of F0, and as a whole the participants had higher F0 variance (or higher pitch dynamism quotient -PDQ-) in DJ questions (mean: 0.214), YN questions (mean: 0.211) and CYN questions (mean: 0.200) than in WH (mean: 0.173) and TAG questions (mean: 0.171). In addition, Figure 12 indicates that speakers of all three proficiency levels had consistently higher F0 variability in questions with penultimate syllable stressed words (mean: 0.200) than those with stress on the final syllable (mean: 0.188). In addition to those features, it is interesting to note that the two Chinese groups produced an overall lower PDQ (CI: 0.168; CA: 0.198) than the L1 native speakers (SN: 0.216), although the factor Proficiency did not show significant main effect in this analysis. Nevertheless, the statistical model revealed a strong interaction between Question type and Proficiency ($F_{(8, 518)}=8.30$, $p<0.001$). As can be seen in Figure 11, compared to the two Chinese groups, the Spanish L1 speakers were higher in PDQ for all question types, in contrast to WH questions in which native speakers produced the lowest PDQ. More precisely,

the post-hoc test indicated that the native group had statistically higher PDQ in DJ questions than the CI (t(2)=-3.317, *p*<0.01) and CA groups (t(2)=-2.593, *p*<0.05), as well as in YN questions when compared to the CI group (t(2)=-3.218, *p*<0.01). No other factors or interactions reached significance on the variable of PDQ.

Finally, in order to examine whether, and to what degree, the F0 span is interdependent on other pitch variables (i.e. mean F0, max F0, min F0 and PDQ), we analyzed the correlation between the span and the rest of the variables that we have taken into account in the pitch range analysis. The statistical results indicated that the 100% and the 80% span were more closely correlated with the maximum F0 (100% span: r=0.84, *p*<0.001; 80% span: r=0.66, *p*<0.001) than with the mean F0 (100% span: r=0.38, *p*<0.001; 80% span: r=0.41, *p*<0.001) or the minimum F0. These results appear to suggest that the more a speaker is able to approximate to the high pitch targets, the wider the entire F0 range of his/her speech. Additionally, there was a strong positive correlation of PDQ with the 100% span (r=0.78, *p*<0.001) and the 80% span (r=0.80, *p*<0.001), as illustrated in the right panel of Figure 13. The high consistency between the two span measures and the PDQ measure also consolidated the results of pitch range variation in our study, despite the unbalanced group size across gender and language group. In general, our findings seem to suggest that the wider the F0 span, the more variable the speech.

Figure 11. Mean PDQ of the three language groups depending on proficiency and question type
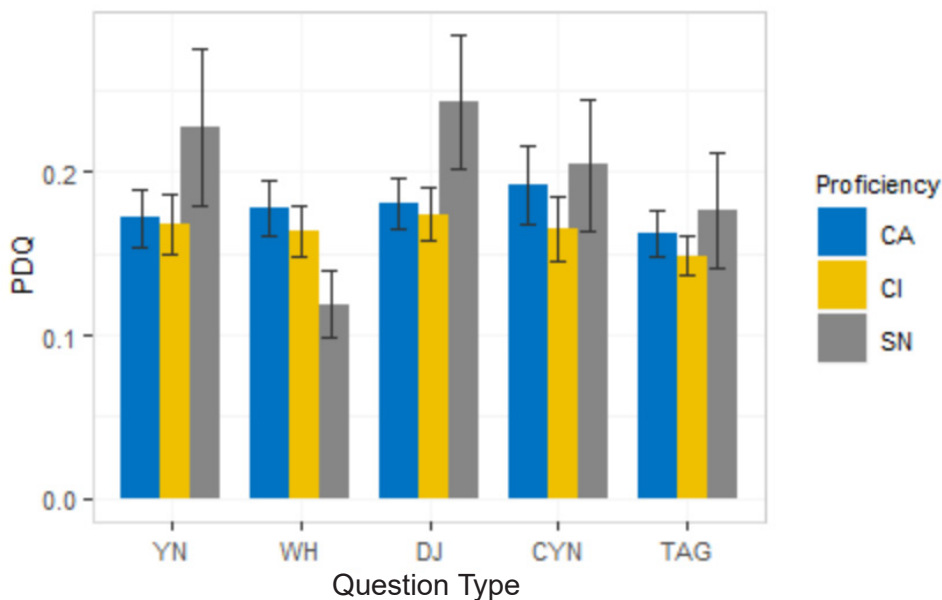
Figure 12. Mean PDQ of the language three groups depending on proficiency and stress position
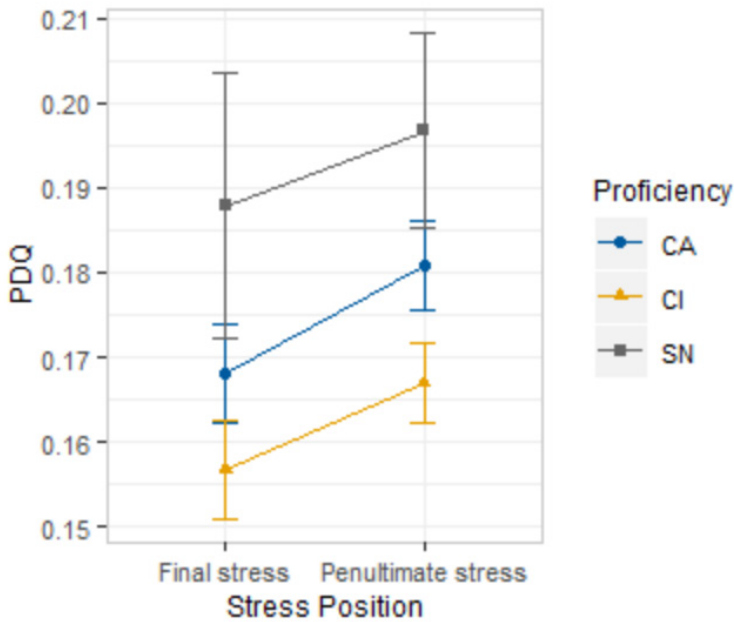
Figure 13. Correlation of the 80% F0 span with the maximum F0 (left) and with the PDQ (right)
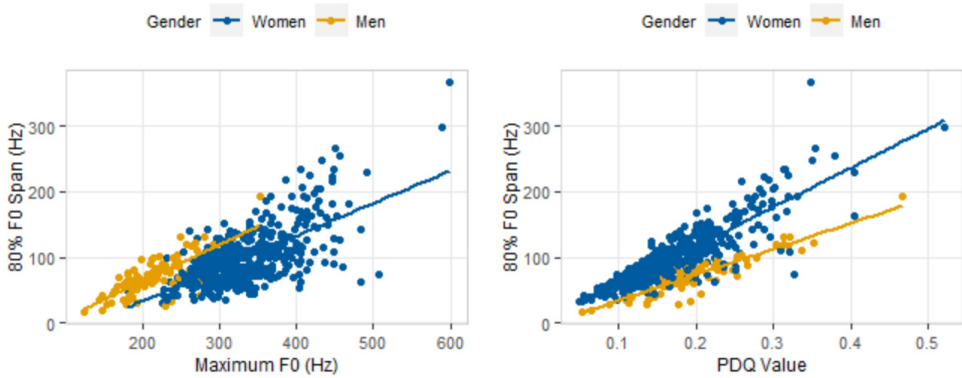
6. Discussion

The present study sets out to investigate the acquisition of L2 Spanish prosody by Chinese L1 speakers. For this purpose, percentages of occurrence of the pitch accents,

boundary tones and nuclear configurations related to different speech acts, as well as various pitch implementation variables have been examined across the three language groups.

The first research question enquired whether the acquisition of intonation contours in an L2 reflects different levels of proficiency, and whether prosodic knowledge from the L1 can be transferred onto L2 pitch movements. In general, it was found that the L2 intonation performance of the phonological contours is dependent on proficiency level, and most of the intonation deviations and success in the L2 could be accounted for under an L1 transfer hypothesis; therefore the answer to the first research question is affirmative. Specifically, the descriptive statistics suggested that compared to the L2 intermediate speakers, L2 advanced learners generally acquired more native-like patterns in producing pitch accents and boundary tones. However, these differences between the two learner groups did not reach statistical significance in our study, probably due to the modest gap of oral proficiency between the Chinese intermediate (B2 level) and advanced group (mostly C1 level). Furthermore, our study found that both the L2 intermediate and the L2 advanced groups produced a significantly higher proportion of high plateaux (H*) and rising accents (L+H*) in the nuclear location than the Spanish L1 speakers, who did not produce any of the aforementioned accent types but showed a statistically higher number of low tones (L*) or falling movements (H+L*) instead. These deviation errors produced by Chinese speakers have been mostly correlated with the negative transfer of their L1. As described before, the lack of a steady low tone in Mandarin's lexical tone inventory and its different stress realization mechanisms (through the expansion of high tone) may be contributing factors in the frequent use of high/rising accents in L2 Spanish intonation. Generally, our findings regarding high/rising tone strategies are consistent with previous studies in L2 speech, which also reported a similar tendency for Chinese learners of other intonation languages such as English and German (Ding et al., 2012; Ji et al., 2009, 2012; Yuan et al., 2019). Moreover, it was surprising that the two Chinese groups, particularly the intermediate group, produced a significantly larger number of high boundaries (H%) and fewer low prosodic tones (L%) than L1 native speakers at the end of the five question types. Such results were contrary to our initial expectations that Chinese learners would use more final falls than rises in L2 Spanish intonation, due to the negative influence of L1 question marking strategies using a particle. A possible explanation for this phenomenon might be that Chinese speakers had at their disposal few intonation contours of the target language and therefore could not assign the correct pattern of different sentence types with specific functional meanings, thus, they tended to overproduce all the questions with the most prototypical final rising intonation. Nevertheless, in general, our results seem to suggest that the performance of L2 speakers in the production of pitch accents and boundary tones can progress

towards native-like shapes by increasing their proficiency level and phonological awareness in the L2.

As for the nuclear configurations, it was found that, excluding the low-rising contours (L* H%), more than half of the learners deviated from the L1 native speakers in the overproduction of high *plateaux* (H* H%) or early rising accents (L+H* H%) in information-seeking *yes-no* questions. Moreover, unlike Spanish L1 speakers who consistently used a high boundary tone to mark the information-seeking *yes-no* question, the final boundaries of Chinese learners could be either high (H%) or low (L%), although low boundaries constituted only a small proportion in their L2 intonation. These results corroborate previous studies on L2 intonation (Ding et al., 2012; Ji et al., 2009, 2012) that have reported similar problems for Chinese L1 speakers of L2 English regarding *yes-no* question patterns. The prevalence of high/rising accents was also observed in the nuclear accents of information-seeking *wh*-questions and confirmation-seeking *yes-no* questions, as a result of the systematic negative transfer from L1 prosody explained above. But in the case of confirmation-seeking *yes-no* questions, most L2 learners have borrowed the intonation inventory of information-seeking *yes-no* questions and only a few of them were able to produce the most salient falling pattern (H+L* L%) for the confirmative speech act. The high comparability in the realization of the two questions types might be attributed to the fact that they share the same syntactic structure in Spanish. However, it is possible that learners faced difficulties in correlating different functional meanings with the specific pitch contour shapes in the L2. Besides, we found that the nuclear patterns of *wh*-questions were variable even in L1 native speech, likely due to individual differences in the interpretation of the situational contexts. Whilst the falling (L* L%) and rising contours (L* H%, with a nuance of politeness and interest) tend to be interpreted as neutral contours for *wh*-questions, the complex rising-falling patterns (L+¡H* L%) in our dataset were biased forms that focused on the reintroduction of topics instead of requesting information. Overall, it seems that the rate of acquisition of target intonation contours is dependent, to some extent, on the systematic similarities between the first prosody and the target prosody, and in this way, we can reasonably explain the great success of Chinese learners in producing L2 disjunctive questions and tag questions as a positive transfer of L1 intonation.

The second research question asked whether the L2 acquisition of pitch implementation details reflects different levels of proficiency, and whether the pitch range values differ among question type, stress position, and gender. The overall results seem to give an affirmative answer to the first part of the question. More precisely, our study found a better pitch performance for the L2 advanced learners than for the L2 intermediate speakers, although the factor "proficiency" did not reach

significance on any of the three pitch measures (except the 80% span measured in St). In addition, the significant interactions of proficiency with other dependent variables show that the L2 prosody acquisition is more complex than previously stated and the proficiency effect was modulated by question type and stress position. Furthermore, the significant main effect of question type for all the pitch measures appears to give a positive answer to the second research question, showing that different sentence types are encoded with different pitch profiles which serve different pragmatic purposes. In other words, the language-specific pitch implementation values were highly dependent on question type. Specifically, we found that the intermediate learners' pitch increased significantly in information-seeking *wh*-questions and confirmation-seeking *yes-no* questions. In addition, it was observed that the two learner groups compressed the span and the F0 variability in all of the question types except for the *wh*-questions in which Chinese informants used wider span and greater F0 variation compared to the L1 native speakers.

As for the pitch differences on stress position, it was interesting to observe that the Chinese learners tended to increase the F0 register, expand the pitch span and exhibit more frequent F0 contour variation in questions where the stress falls on the penultimate-syllable of the word than those with stress on the final syllable of the word, whereas the native speakers (on the contrary) produced a wider span and higher register for the oxytone words. The factor which accounts for these differences between L1 and L2 is mainly correlated with learners' unfamiliarity and lack of knowledge of rules in producing the final-syllable stressed words, particularly when they appear in the sentence-final position that requires compressed pitch movements. In such circumstances, learners may pay more attention to the pronunciation of words and reduce the use of F0 strategies in the L2 (Mennen, Schaeffler, & Dickie, 2014; Peters, 2019; Yuan et al., 2018; Zimmerer et al., 2014). Finally, with regard to gender effects, our results are congruent with previous findings which report that women use significantly higher pitch level and greater pitch variation than men. Hence, the answer to the latter part of the second research question is also affirmative. According to Ohala (1994), Van Bezooijen (1995), Urbani (2012), Mennen et al. (2014) and Peters (2019)weak, dependent, modest, pitch differences across gender are not merely driven by physiological and anatomical factors, but are also closely linked to some sociocultural aspects such as social status and gender roles. Moreover, it is assumed that the greater F0 variation of female speakers found in our study might be correlated with the finding that women tend to express many emotions more frequently than men, except for pride and power (Brebner, 2003). However, further research on the current topic are required to provide direct and strong evidence for our speculation. Finally, in the data obtained for our study, we also found a strong positive correlation between the F0 span and the maximum F0 as well as the F0 variability. More precisely,

our study appears to suggest that the wider the entire F0 span, the higher the F0 pitch targets and the more variable the F0 speech.

The third research question enquired whether our pitch implementation findings point towards a universal developmental path during the L2 learning process, regardless of the language combinations under study. In the present study, we found that compared to Spanish L1 speakers, Chinese learners tended to use higher pitch level (specifically higher minimum and mean F0), narrower span and less variable pitch in the L2. These results are in agreement with other preliminary studies on L2 speech which also documented a compressed pitch pattern for L2 speakers of different language backgrounds, and reported a similar trend towards improvement as their L2 experience increased (Busà & Urbani, 2011; Mennen, Schaeffler, & Docherty, 2009; Mennen, 1998; Peters, 2019; Shi, Zhang, & Xie, 2014; Ullakonoja, 2007; Urbani, 2012; Yuan et al., 2018; Zimmerer et al., 2014). As for the differences of F0 height, the higher pitch level observed in L2 intonation has been explained as a result of an increased cognitive effort when speaking a non-dominant language (Zimmerer et al., 2014). However, in our study, we speculate that the L1 pitch characteristics may also play a role in the F0 rise of L2. To test this hypothesis, further investigations comparing L1 Chinese and L1 Spanish are needed.

The final research question examined whether the acquisition of the phonological and phonetic patterns of an L2 reflects different levels of difficulty corresponding to pragmatically different question types. On the one hand, it is interesting to note that the probability of success with regard to the production of L2 nuclear contours was different across the five question types. Therefore, the answer to the last research question must be affirmative in the phonological dimension. Specifically, our results seem to suggest a hierarchy of difficulty in the intonation learning of different question patterns of Spanish by Chinese L1 speakers, whereby the confirmation-seeking *yes-no* question was the most difficult pattern, followed by the information-seeking *yes-no/ wh* question, then the disjunctive question and finally the confirmation-seeking tag question. Along the same lines, Cortés Moreno (1999, 2004), Liu (2005), Mennen (2015) and Yuan et al. (2019) have also reported that different sentence types may imply different degrees of difficulty during the learning process of L2 phonological contours. On the other hand, the varied pitch performance among the five question types appears to extend previous hypotheses on L2 intonation (Cortés Moreno, 1999, 2004; Liu, 2005; Yuan et al., 2019) from the phonological to the phonetic dimension, suggesting that there is also a ranking of difficulty in implementing the pitch values of different question patterns in L2. The relatively faster rate of acquisition of the two intonational aspects (phonological and phonetic) has been partially correlated with the typological closeness between the first language and the target language.

Moreover, this has also been explained as a result of the "perceptual salience" of some intonational movements in the target language (Yuan et al., 2019).

7. Conclusion

All things considered, the present study has shown that the L2 intonation learning process is more complex than previously stated, whereby the phonological and phonetic dimensions may develop in a non-parallel way. Specifically, it is worth noting that learners who have successfully acquired the target-like intonation contours may still deviate from native speakers in the pitch implementation details of the L2 prosodic system. However, as their L2 proficiency increases, they are capable of approximating to the target language settings both in the systematic dimension and in some realizational aspects. This finding of non-uniform development paths for L2 phonetic and phonological acquisition has also been observed in other L1-L2 language pairs and seems to be a universal feature that occurs throughout the course of foreign language learning (Ding et al., 2012; Graham & Post, 2018; Mennen & Leeuw, 2014; Mennen, 2015). In addition, the distinct pitch performance in L2 appears to suggest a differing degree of difficulty in acquiring Spanish intonation depending on sentence type, stress position and gender. Beyond this, our study extends previous hypotheses by proposing a progression in difficulty levels from the phonological to the phonetic dimension, suggesting that this difficulty ranking exists not merely when acquiring the L2 phonological contours, but also when implementing the pitch values of different sentence types.

From a teaching perspective, our study sheds light on L2 prosody learning, particularly on the L2 acquisition of Spanish intonation by Chinese learners which, despite its crucial importance, has not been addressed in many publications. Overall, it is proposed that there should be special training methods based on specific tasks to help reduce learners' foreign accents, as they showed distinct performances for L2 intonation patterns, according to the similarities and dissimilarities between the first and target language. Furthermore, our results suggest that the training program should not only include the phonological knowledge of target intonation contours but also, and perhaps more importantly, should allow learners to interact with language-appropriate contexts and to produce pitch implementation details in a native-like way. At this point, a growing number of recent investigations are devoted to the development of intonation teaching techniques. For instance, pitch gestures have been reported to benefit L2 prosody learning (Baills, 2016; Bernardis & Gentilucci, 2006; Gullberg, 2006; Jia & Wang, 2013a, 2013b; Morett & Chang, 2015), particularly in acquiring the low nuclear accents, which constitute the most difficult patterns for Chinese L2 learners. Other approaches, such as music training activities and speech visualization

tools, can also help learners to progress in L2 speech production. Nevertheless, considering that Spanish intonation has different levels of difficulty depending on sentence type, any pedagogical proposal should include scaffolding techniques in order to ease the way into the phonetic implementation.

Finally, some limitations should be noted in the current study. One such limitation is the lack of more precise criteria on the sample profile. Specifically, due to the dramatic reduction of possible samples, we did not specifically control for Chinese subjects' origin, age of L2 acquisition, or length of exposure to the target-language environment. Moreover, the sample sizes of male and female speakers were unequal across the three language groups. However, the effect of gender imbalance on the pitch range values was minimized by introducing the PDQ measure, which can effectively normalize the F0 variation data and make the LTD results more robust. Another limitation is related to the elicitation of the corpus. Due to the ambiguity of the situational contexts designed for the information-seeking *wh*-question, there were some unexpected nuclear pitch patterns in the L1 production, for instance, L+¡H* L%, which has been described as a form of the reintroduction of a previous topic rather than a request for new information. Finally, although our study found that there were some general pitch range deviations for L2 speakers, we did not discover exactly how the F0 range was realized depending on the syntactic position of the phrase and in which positions the learners deviated most from the L1 pitch patterns. Therefore, further research could take into account the effect of syntactic position to examine whether the increased pitch level and the compressed pitch span occur along with the overall L2 utterances or whether they are position-sensitive (e.g. increasing only on the low pitch targets while high tones remain basically unchanged).

8. Acknowledgments

9. References

Abu-Al-Makarem, A. & Petrosino, L. (2007) Reading and spontaneous speaking fundamental frequency of young Arabic men for Arabic and English languages: A comparative study. *Perceptual and Motor Skills*, 105(2), 572–580.

Andreeva, B., Demenko, G., Wolska, M., Möbius, B., Zimmerer, F., Jügler, J., Magdalena, O.P. & Trouvain, J. (2014) Comparison of pitch range and pitch variation in Slavic and Germanic languages. In N. Campbell, D. Gibbon & D. Hirst (eds) *Proceedings of the 7th International Conference on Speech Prosody*. (pp. 776–780). Dublin, Ireland: Trinity College Dublin.

Atoye, R. O. (2005) Non-native perception and interpretation of English intonation. *Nordic Journal of African Studies*, 14(1), 26-42.

Backman, N. (1979) Intonation Errors in Second-Language Pronunciation of Eight Spanish-Speaking Adults Learning English. *Interlanguage Studies Bulletin*, 4(2), 239–265.

Baills, F. (2016). *Producing pitch gestures enhances the acquisition of Mandarin Chinese tones*. Unpublished MA thesis. Departament of Translation and Language Science, Universitat de Ponpeu Fabra.

Barto, K. A. (2015). *Mandarin Speakers' Intonation in their L2 English*. Unpublished doctoral dissertation. Program in Second Language Acquisition and Teaching, University of Arizona.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48.

Beckman, M. E. & Venditti, J. J. (2011) Intonation. In J. Goldsmith, J. Riggle & A. C. L. Yu (eds) *The handbook of phonological theory*. (pp. 485–532). Oxford, UK: Blackwell.

Benesty, J., Chen, J., Huang, Y. & Cohen, I. (2009) Pearson correlation coefficient. In I. Cohen, Y. Huang, J. Chen & J. Benesty (eds) *Noise reduction in speech processing*. (pp. 1–4) Heidelberg, Germany: Springer.

Bernardis, P. & Gentilucci, M. (2006) Speech and gesture share the same communication system. *Neuropsychologia*, 44(2), 178–190.

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

Bowen, J. D. (1956) A comparison of the intonation patterns of English and Spanish. *Hispania*, 39(1), 30–35.

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 W.-S. Lee & E. Zee (eds) *Proceedings of the 17th International*

Congress of Phonetic Sciences. (pp. 380–383). Hong Kong, China: City University of Hong Kong.

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 short-term L2 proficiency in young Danish learners of English. *Vigo International Journal of Applied Linguistics*, 17, 57–86.

Chao, Y. R. (1968) *A Grammar of Spoken Chinese*. Beijing: The Commercial Press.

Chen, G. T. (1972). *A comparative study of pitch range of native speakers of Midwestern English and Mandarin Chinese: An acoustic study*. Unpublished doctoral dissertation. Department of Philosophy, University of Wisconsin-Madison.

Chen, S. H. (2005) The effects of tones on speaking frequency and intensity ranges in Mandarin and Min dialects. *The Journal of the Acoustical Society of America*, 117(5), 3225–3230.

Chen, Y. & Gussenhoven, C. (2008) Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics*, 36(4), 724–746.

Cook, V. (2012) The native speaker and Multicompetence. In P. Robinson (ed) *Routledge Encyclopedia of Second Language Acquisition*. (pp. 447-450). London, UK: Routledge.

___. (2016) Premises of multicompetence. In V. Cook & W. Li (eds) *The Cambridge Handbook of Linguistic Multicompetence*. (pp. 1-25). Cambridge: Cambridge University Press.

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.

___. (1999) *Adquisición de la entonación española por parte de hablantes nativos de chino*. Unpublished doctoral dissertation. Departamento de Didáctica de la Lengua y la Literatura, Universitat de Barcelona.

___. (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.

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

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

Cruttenden, A. (1997) *Intonation*. Cambridge: Cambridge University Press.

Cuenca, M. J. (1997) Form-use mappings for tag questions. *Amsterdam Studies in the Theory and History of Linguistic Science*, 4, 3–20.

Ding, H., Hoffmann, R. & Hirst, D. (2016) Prosodic transfer: A comparison study of f0 patterns in L2 English by Chinese speakers. *Proceedings of the 8th International Conference on Speech Prosody* (pp. 756–760). Boston, USA: Boston University.

Ding, H., Jokisch, O. & Hoffmann, R. (2012) A phonetic investigation of intonational foreign accent in Mandarin Chinese learners of German. In Q. Ma, H. Ding & D. Hirst (eds) *Proceedings of the 6th International Conference on Speech Prosody*. (pp. 374–377). Shanghai, China: Tongji University Press.

Elvira-García, W., Roseano, P., Fernández-Planas, A. M. & Martínez-Celdrán, E. (2016) A tool for automatic transcription of intonation: Eti_ToBI a ToBI transcriber for Spanish and Catalan. *Language Resources and Evaluation*, 50(4), 767–792.

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). München, Germany: Lincom Europa.

Estebas-Villaplana, E. (2008) The evaluation of intonation: pitch range differences in English and in Spanish. In G. Thompson & L. Alba-Juez (eds) *Evaluation in Context*. (pp. 285–301) Amsterdam, The Netherlands: John Benjamins.

Face, T. L. & Prieto, P. (2007) Rising accents in Castilian Spanish: A revision of Sp_ToBI. *Journal of Portuguese Linguistics*, 6(1), 117–146.

Félix-Brasdefer, J. C. (2010) Data collection methods in speech act performance: DCTs, role plays, and verbal reports. *Speech Act Performance: Theoretical, Empirical and Methodological Issues*, 26, 41-56.

Font-Rotchés, D. (2005). *L'entonació del català. Patrons melòdics, tonemes i marges de dispersió.* Unpublished doctoral dissertation. Department of Didactics and Literature, Universitat de Barcelona.

Frota, S. & Prieto, P. (2015) *Intonation in Romance*. Oxford: Oxford University Press.

Girding, E., Zhang, J. & Svantesson, J.-O. (1983) A generative model for tone and intonation in Standard Chinese. *Working Papers (Lund University: Department of Linguistics)*, 25, 53-65.

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

Gullberg, M. (2006) Some reasons for studying gesture and second language acquisition (Hommage à Adam Kendon). *IRAL-International Review of Applied Linguistics in Language Teaching*, 44(2), 103–124.

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

Gussenhoven, C. & Chen, A. (2000) Universal and language-specific effects in the perception of question intonation. In B. Yuan, T. Huang & X. Tang (eds) *Proceedings of the 6th International Conference on Spoken Language Processing*. (pp. 91–94). Beijing, China: China Military Friendship Publish.

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

Hewings, M. (1995) The English Intonation of Native Speakers and Indonesian Learners: a Coniparative Study. *RELC Journal*, 26(1), 27–46.

Hong, W. (2012) *An experimental study of Chinese students' English intonation pattern*. Unpublished doctoral dissertation. Department of Engllish Philosophy, Nankai University.

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.

Hualde, J. I. (2003) El modelo métrico y autosegmental. In P. Prieto (coord) *Teorías de La Entonación*. (pp. 155–184). Barcelona, Spain: Ariel.

Hualde, J. I. & Prieto, P. (2015) Intonational variation in Spanish: European and American varieties. In S. Frota & P. Prieto (eds) *Intonation in Romance*. (pp. 350–391). Oxford, UK: Oxford University Press.

Hyman, L. M. (2006) Word-prosodic typology. *Phonology*, 23(2), 225–257.

___. (2012) In defense of prosodic typology: A response to Beckman and Venditti. *Linguistic Typology*, 16(3), 341–385.

Ji, X. (2010) *Acquisition of intonation by Chinese EFL learners–An empirical study based on experimental phonetics*. Unpublished MA thesis. Department of English Linguistics, Zhejiang University.

Ji, X., Wang, X. & Li, A. (2009) Intonation patterns of yes-no questions for Chinese EFL learners. *Proceedings of the 2009 Oriental COCOSDA International Conference on Speech Database and Assessments*. (pp. 88–93). Urumqi, China: IEEE.

Ji, X., Xia, W., Yuan, J. & Li, A. (2012) Intonation patterns of *wh*-questions from Chinese EFL learners. In Q. Ma, H. Ding & D. Hirst (eds.) *Proceedings of the*

6th International Conference on Speech Prosody. (pp. 741–744). Shanghai, China: Tongji University Press.

Jia, L. & Wang, J. Q. (2013a). On the effects of visual processing on tone production by English-speaking learners of Chinese. *TCSOL Studies*, 52(4), 30–34.

___. (2013b.) The effects of visual processing on tone perception by native English-speaking learners of Chinese. *Chinese Teaching in the World*, 27, 548–557.

Jia, Y., Xiong, Z. & Li, A. (2006) The Effect of Focal Accents upon Sentential Pitch in Standard Chinese. *Report of Phonetic Research*, 61-68. Beijing, China.

Jilka, M. (2000) Testing the contribution of prosody to the perception of foreign accent. *New Sounds*, 4, 199–207.

Juan-Garau, M. & Pérez-Vidal, C. (2007) The effect of context and contact on oral performance in students who go on a stay abroad. *Vigo International Journal of Applied Linguistics*, 4, 117–134.

Juffs, A. (1990) Tone, syllable structure and interlanguage phonology: Chinese learners' stress errors. *IRAL-International Review of Applied Linguistics in Language Teaching*, 28(2), 99–118.

Jun, S.-A. (ed) (2006) *Prosodic typology: The phonology of intonation and phrasing.* Oxford: Oxford University Press.

Keating, P. & Kuo, G. (2012) Comparison of speaking fundamental frequency in English and Mandarin. *The 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.

Kochanski, G. & Shih, C. (2003) Prosody modeling with soft templates. *Speech Communication*, 39(3–4), 311–352.

Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. (2017) lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1-27.

Ladd, D. R. (1996) *Intonational phonology.* Cambridge: Cambridge University Press.

___. (2001) Intonational universals and intonational typology. In M. Haspelmath, E. König, W. Oesterreicheran & W. Raible (eds) *Language Typology and Language Universals: An International Handbook.* (pp. 1-18). Berlin, Germany: De Gruyter.

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

Ling, F. (2003). *A Study on the acoustic correlates of emphasized Tone 3 in Mandarin Chinese*. Unpublished doctoral dissertation. Deparament of linguistics, Peking University.

Liu, F. & Xu, Y. (2005) Parallel encoding of focus and interrogative meaning in Mandarin intonation. *Phonetica*, 62(2-4), 70-87.

Liu, Y. H. (2005) La entonación del español hablado por taiwaneses. *Biblioteca Phonica*, 2, 1-807.

Mackey, W.F. (2000) The description of bilingualism. In L. Wei (ed) *The Bilingualism Reader*. (pp.26-54). Oxford, New York: Routledge.

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

May, S. (ed) (2013) *The multilingual turn: Implications for SLA, TESOL, and bilingual education*. New York: Routledge.

McGory, J. T. (1997). *Acquisition of intonational prominence in English by Seoul Korean and Mandarin Chinese speakers*. Unpublished doctoral dissertation. Deparament of Speech and Hearing Sciences,The Ohio State University.

Mennen, I., Schaeffler, F. & Docherty, G. (2007) Pitching it differently: a comparison of the pitch ranges of German and English speakers. In J. Trouvain & W. J. Barry (eds) *Proceedings of the 16th International Congress of Phonetic Sciences*. (pp. 1769-1772). Saarbrücken, Germany: Pirrot GmbH.

Mennen, I. & Leeuw, E. De. (2014) Beyond segments. *Studies in Second Language Acquisition*, 36(2), 183-194.

Mennen, I. (1998) Can language learners ever acquire the intonation of a second language? In *ESCA Workshop on Speech Technology in Language Learning (STiLL 98)*. (pp. 17-20). Marholmen, Sweden: Department of Speech, Music and Hearing, KTH.

___ (2015) Beyond segments: towards an L2 intonation learning theory (LILt). In E. Delais-Roussarie, M. Avanzi & S. Herment (eds) *Prosody and Language in Contact: L2 Acquisition, Attrition and Languages in Multilingual Situations*. (pp. 171-188). Berlin, Germany: Springer.

Mennen, I., Chen, A. & Karlsson, F. (2010) Characterising the internal structure of learner intonation and its development over time. In K. Dziubalska-Kołaczyk, M. Wrembel & M. Kul (eds) *Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech*. (pp. 319-324). Poznań, Poland: Adam Mickiewicz University.

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

Mennen, I. (2007) Phonological and phonetic influences in non-native intonation. In J. Trouvain & U. Gut (eds) *Non-Native Prosody: Phonetic Description and Teaching Practice*. (pp. 53–76). Berlin, Germany: De Gruyter Mouton.

Morett, L. M. & Chang, L.-Y. (2015) Emphasising sound and meaning: Pitch gestures enhance Mandarin lexical tone acquisition. *Language, Cognition and Neuroscience*, 30(3), 347–353.

Ng, M. L., Hsueh, G., Leung, S. & Shing, C. (2010) Voice pitch characteristics of Cantonese and English produced by Cantonese-English bilingual children. *International Journal of Speech-Language Pathology*, 12(3), 230–236.

Niebuhr, O. & Skarnitzl, R. (2019) Measuring a speaker's acoustic correlates of pitch–but which? A contrastive analysis based on perceived speaker charisma. In S. Calhoun, P. Escudero, M. Tabain & P. Warren (eds) *Proceedings of the 19th International Congress of Phonetic Sciences*. (pp. 1774-1778). Melbourne, Australia: Australasian Speech Science and Technology Association Inc.

Nolan, F. (2003) Intonational equivalence: an experimental evaluation of pitch scales. In M. J. Solé, D. Recasens & J. Romero (eds) *Proceedings of the 15th International Congress of Phonetic Sciences*. (Vol. 771, pp. 2–5). Barcelona, Spain.

___. (2006) Intonation. In B. Aarts & A. McMahon (eds) *The Handbook of English Linguistics*. (pp. 433–457). Hoboken, USA: Blackwell Publishing Ltd.

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

___. (1994) The frequency code underlies the sound-symbolic use of voice pitch. In L. Hinton, J. Nichols & J. J. Ohala (eds) *Sound symbolism*. (pp. 1337-1340). Cambridge, UK: Cambridge University Press.

Ortega, L. (2010) The Bilingual Turn in SLA. Plenary delivered at the *Annual Conference of the American Association for Applied Linguistics*. Atlanta, GA, March 6-9.

___. (2011) New trends in SLA research: Theories, methods, ethics. *Invited lecture at National Tsing Hua University*. Taiwan, June 8.

___. (2014) Ways forward for a bi/multilingual turn in SLA. In Stephen M. (ed) *The multilingual turn: Implications for SLA, TESOL and bilingual education*. (pp. 32-53). New York: Routledge.

Pandey, A. (2016). The multilingual turn: Implications for SLA, TESOL, and bilingual education. Stephen M. (ed). 2014. New York: Routledge. *World Englishes*, 35(3), 475-477.

Payà, M. & Vanrell, M. del M. (2005) Yes-no questions and echo questions intonation in Majorcan and Minorcan Catalan: A cross-dialectal comparison. In M. Zdrojkowski (ed) *Phonetics and Phonology in Iberia (PaPI).* (pp. 20–21). Barcelona, Spain.

Peters, J. (2019) Fluency and speaking fundamental frequency in bilingual speakers of High and Low German. In S. Calhoun, P. Escudero, M. Tabain & P. Warren (eds) *Proceedings of the 19th International Congress of Phonetic Sciences.* (pp. 1-5). Melbourne, Australia: Australasian Speech Science and Technology Association Inc..

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*, 3, 147–179.

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation.* Unpublished doctoral dissertation. Department of Linguistics and Philosophy, Massachusetts Institute of Technology.

Prieto, P. & Roseano, P. (2019) Current issues and challenges in Spanish intonational research. In S. Colina & M.-G. Fernando (eds) *The Routledge Handbook of Spanish Phonology.* (pp. 222–233). UK: Taylor & Francis.

Rendall, D., Vokey, J. R. & Nemeth, C. (2007) Lifting the curtain on the Wizard of Oz: Biased voice-based impressions of speaker size. *Journal of Experimental Psychology: Human Perception and Performance*, 33(5), 1208–1219.

Roseano, P., González, M., Borràs-Comes, J. & Prieto, P. (2016) Communicating epistemic stance: How speech and gesture patterns reflect epistemicity and evidentiality. *Discourse Processes*, 53(3), 135–174.

Schack, K. (2000) Comparison of intonation patterns in Mandarin and English for a particular Speaker. In K. M. Crosswhite & J. McDonough (eds) *Working Papers in the Language Science.* (pp. 24-55). Rochester, USA: University of Rochester.

S.J.Eady. (1982) Differences in the f0 patterns of speech: Tone language versus stress language. *Language and Speech*, 25, 29–42.

Shao, P., Jia, Y. & Li, A. (2011) Intonation patterns of exclamations of Chinese EFL learners from Zhejiang. *Proceedings of the 2011 International Conference on Speech Database and Assessments (Oriental COCOSDA).* (pp. 74–78). Hsinchu City, Taiwan: IEEE.

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 M. Dong, J. Tao, H. Li, T. F. Zheng & Y. Lu (eds) *Proceedings of the 9th International Symposium on Chinese Spoken Language Processing.* (pp. 241–244). Singapore: IEEE.

Shih, C. (1988) Tone and intonation in Mandarin. *Working Papers, Cornell Phonetics Laboratory*, 3, 83–109.

J't Hart, J., Collier, R. & Cohen, A. (1990) *A Perceptual Study of Intonation: An Experimental Phonetic Approach to Speech Melody*. Cambridge: Cambridge University Press.

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). Amsterdam, The Netherlands: Elsevier Science B.V.

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

Team, R. C. (2016) R: *A language and environment for statistical computing* [Computer software manual]. Vienna, Austria.

Todaka, Y. (1990). *An error analysis of Japanese students' intonation and its pedagogical applications*. Unpublished MA thesis. Department of Lingsuitics, University of California.

Ullakonoja, R. (2007) Comparison of pitch range in Finnish (L1) and Russian (L2). In J. Trouvain & W. J. Barry (eds) *Proceedings of the 16th International Congress of Phonetic Sciences*. (pp. 1701–1704). Saarbrücken, Germany: Pirrot GmbH.

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 Second Language prosody*. (pp. 79–83). Padova, Italia: Università di Padova

Valdés, G., Poza, L. & Brooks, M. D. (2015) Language acquisition in bilingual education. In Wayne E. W., Sovicheth B. & Ofelia G. (eds) *The handbook of bilingual and multilingual education*. (pp. 56-74). Malden, MA: Wiley-Blackwell.

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

Van Dommelen, W. A. & Moxness, B. H. (1995) Acoustic parameters in speaker height and weight identification: sex-specific behaviour. *Language and Speech*, 38(3), 267–287.

Vanrell, M. del M., Mascaró, I., Torres-Tamarit, F. & Prieto, P. (2010) When intonation plays the main character : information-vs.confirmation-seeking questions in Majorcan Catalan. *Proceedings of the 15th International Conference on Speech Prosody*. (pp. 4–7). Chicago, USA: S. Karger AG.

Vilaplana, E. E. (2008) Modelling final declarative intonation in English and Spanish. In T. Gibert & L. Alba (eds) *Estudios de Filología Inglesa: Homenaje a La Dra. Asunción Alba Pelayo*. (pp. 109–129). Madrid, Spain: Ediciones UNED.

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.

Wennerstrom, A. (1998) Intonation as cohesion in academic discourse: A study of Chinese speakers of English. *Studies in Second Language Acquisition*, 20(1), 1–25.

Wu, Z. (1982) Rules of intonation in Standard Chinese. *Preperation for Working Group on Intonation, 8th International Congress of Linguists.* (pp. 95–108). Tokyo, Japan.

Xu, H. (2009). *A survey study of Chinese EFL learners' acquisition of English intonation: A functional perspective.* Unpublished MA thesis. Department of English Philosophy, Jiangsu University.

___. (2005) Speech melody as articulatorily implemented communicative functions. *Speech Communication*, 46(3–4), 220–251.

___. (2017) Intonation in Chinese. In W. Behr, Y. Gu, Z. Handel, C.-T. J. Huang & J. Myers (eds) *Encyclopedia of Chinese language and linguistics.* (pp. 458–466). Boston: Brill.

Xu, Y. & Wang, Q. E. (2001) Pitch targets and their realization: Evidence from Mandarin Chinese. *Speech Communication*, 33(4), 319–337.

Yamazawa, H. & Hollien, H. (1992) Speaking fundamental frequency patterns of Japanese women. *Phonetica*, 49(2), 128–140.

Yip, M. (2002) *Tone.* Cambridge: Cambridge University Press.

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. (2004). *Intonation in Mandarin Chinese: Acoustics, perception, and computational modeling.* Unpublished doctoral dissertation. Department of lingusitics, Cornell University.

___. (2006) Mechanisms of question intonation in Mandarin. In B. Ma, Huo Q., E.-S. Cheng & H. LI (eds) *Proceedings of the 5th International Symposium on Chinese Spoken Languag.* (Vol. 4274, pp. 19–30). Berlin, Germany: Springer.

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. *Proceedings of the 19th Annual Conference of the International Speech Communication Association.* (pp. 2593–2597). Hyderabad, India: Curran Associates, Inc.

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

Yuan, J. & Shih, C. (2004) Confusability of Chinese Intonation. In B. Bernardl & M. Isabelle (eds) *Proceedings of the 2nd International Conference on Speech Prosody*. (pp. 1–4). Nara, Japan.

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 N. Campbell, D. Gibbon & D. Hirst (eds) *Proceedings of the 7th International Conference on Speech Prosody*. (pp. 1037–1041). Dublin, Ireland: Trinity College Dublin.

Appendix A

Biographical information for L2 learners in the study, with Mandarin Chinese as the L1.

Code	Gender	Age at test	Age of acquisition	Month of residence in Spain	L2 proficiency	China origin	L2 dialect
ci01	F	22	18	10	B2	Henan	PS
ci02	F	22	18	10	B2	Sichuan	PS
ci03	M	26	21	12	B2 (SJ)	Shandong	PS
ci04	F	22	18	12	B2	Liaoning	PS
ci05	M	23	19	13	B2	Jiangsu	PS
ci06	F	23	19	14	B2	Jiangsu	PS
ci07	F	24	21	20	B2 (SJ)	Liaoning	PS
ci08	F	22	18	22	B2	Jiangsu	PS
ci09	M	31	24	22	B2	Shandong	PS
ci10	F	22	18	22	B2	Zhejiang	PS
ci11	F	23	17	25	B2	Anhui	PS
ci12	M	24	18	25	B2	Jiangxi	PS
ci13	M	24	18	25	B2	Shanxi	PS
ci14	M	21	19	32	B2 (SJ)	Guangdong	PS
ci15	F	23	19	40	B2 (SJ)	Ningxia	PS
ci16	F	23	18	60	B2 (SJ)	Shanghai	PS
ca17	F	24	18	1	C1	Shandong	PS

ca18	F	22	19	2	C1	Gansu	PS
ca19	F	24	20	2	C1	Jilin	PS
ca20	F	23	19	3	C1	Neimenggu	PS
ca21	F	22	18	7	C1	Beijing	PS
ca22	F	22	18	8	C1	Shanghai	PS
ca23	F	24	19	10	C1	Shanxi	PS
ca24	F	22	18	11	C1	Heilongjiang	PS
ca25	F	26	24	24	C1	Guangxi	PS
ca26	F	25	18	25	C1	Heilongjiang	PS
ca27	F	30	19	25	C2	Liaoning	PS
ca28	F	24	22	25	C1	Zhejiang	PS
ca29	F	29	18	36	C1	Henan	PS
ca30	F	24	18	40	C1	Heilongjiang	PS
ca31	F	29	19	40	C1	Anhui	PS
ca32	F	26	12	48	C2	Tianjin	PS

Note: Proficiency (SJ = self-judgement by L2 learners); L2 dialect (PS = Peninsular Spanish).

Appendix B

Test items

(1)	Sample contexts for information-seeking *yes-no* question: Entras en una frutería donde no has estado nunca y le preguntas al dependiente si tiene mandarinas. *¿Tiene mandarinas?*
(2)	Sample contexts for information-seeking *wh*-question: Has quedado con dos buenas amigas para hacer compras esta tarde. Pero al llegar la hora, sólo se presenta una chica y no ves a la otra amiga, Marina. Pregúntale dónde está Marina. *¿Dónde está Marina?*
(3)	Sample contexts for disjuntive question: Has invitado a buen amigo a tu piso para una cena. Después de acabar los platos principales, le preguntas si quiere tarta o helado de postre. *¿Quieres tarta o helado?*
(4)	Sample contexts for confirmation-seeking *yes-no* question: Estás hablando con tu compañero de piso y ves que está cubierto con dos mantas. Tú infieres que tu compañero de piso debe de tener frío y le preguntas si es así. *¿Tienes frío?*
(5)	Sample contexts for confirmation-seeking tag question: Vas a una discoteca con un amigo y ves que baila muy bien, así que supones que practica baile a menudo (mucho) y le preguntas si es así. *Bailas a menudo, ¿verdad?*

Other test items

Stress position	Penultimate stressed syllable	Final stressed syllable
	¿Tiene <u>cam</u>bio?	¿Tiene me<u>lón</u>?
	¿Cuándo <u>vie</u>ne?	¿Dónde está Ma<u>nuel</u>?
	¿Viene el <u>lu</u>nes o el <u>mar</u>tes?	¿Quieres me<u>lón</u> o melo<u>co</u>tón?
	¿Tienes <u>ham</u>bre?	¿Tienes ca<u>lor</u>?
	Sales al mer<u>ca</u>do, ¿no?	No te encuentras <u>bien</u>, ¿verdad?