>>HABLA, VOZ Y SONIDO >>SPEECH, VOICE & SOUND >>PARLA, VEU I SO

NUCLEAR CONTOURS OF SPANISH STATEMENTS AND QUESTIONS PRODUCED BY MANDARIN SPEAKERS WITH ADVANCED SPANISH PROFICIENCY

Peng Li
University of Oslo
(Norway)
peng.li@iln.uio.no
https://orcid.org/0000-0003-4488-3514

Xiaotong Xi
Shandong University of Finance and Economics
(China)
xiaotong.xi@sdufe.edu.cn
https://orcid.org/0000-0002-9306-2942

Received: 29/4/2024. Accepted: 6/8/2024. Published: 16/09/2024

Recommended citation:

Li, P., & Xi, X. (2024). Nuclear contours of Spanish statements and questions produced by mandarin speakers with advanced Spanish proficiency. *Phonica*, 20, 1-26. https://doi.org/10.1344/phonica2024.20.3

Abstract

This study dynamically models the nuclear contours of L2 Spanish sentences produced by 16 Mandarin-speaking learners and 9 Spanish natives, using a speech corpus obtained from a discourse completion task. The target sentences included statements (broad, categorical, and corrective focus) and yes/no questions (information-seeking, confirmation-seeking, and tag questions). Our results indicate that Chinese students (a) could not produce correct nuclear configurations to differentiate between the categorical statement and corrective focus; (b) produced a significantly higher pitch for nuclear pitch accent and lower pitch for unstressed syllables compared to Spanish natives; and (c) may have produced incorrect boundary tones influenced by the lexical stress positions of the nuclear word. This study contributes empirical evidence to the theory of L2 prosodic learning and highlights the importance of fine-grained phonetic details beyond phonological (dis)similarities between learners' L1 and L2 prosody. Furthermore, the observed difficulties in L2 prosody among even experienced learners highlight the need for proper prosodic training paradigms in teaching practice.

Keywords: Prosody; L2 Spanish; L1 Chinese speakers; Nuclear configuration; Second language

Contorns nuclears d'oracions declaratives i interrogatives en castellà produïts per parlants de mandarí amb una competència avançada en castellà (Català)

Resum: En aquest estudi, es modelen dinàmicament els contorns nuclears d'oracions en castellà produïdes per 16 estudiants parlants de mandarí i per 9 parlants nadius de castellà, utilitzant un corpus de parla obtingut mitjançant una *discourse completion task*. Les oracions investigades inclouen declaratives (focus ampli, categòric i correctiu) i interrogatives absolutes (informatives, confirmatòries i preguntes amb partícula confirmatòria). Els nostres resultats indiquen que els estudiants xinesos (a) no aconsegueixen produir les configuracions nuclears correctes per diferenciar les oracions declaratives categòriques de les de focus correctiu; (b) pronuncien significativament accents tonals més alts en posició nuclear i accents tonals més baixos en les síl·labes àtones en comparació amb els parlants nadius de castellà; i (c) fan tons de frontera incorrectes degut a la posició de l'accent lèxic a les paraules nuclears. Aquest estudi aporta evidències empíriques a la teoria de l'aprenentatge prosòdic de L2 i remarca la importància dels detalls fonètics per sobre de les similituds/diferències fonològiques entre parlants de L1 i L2. A més a més, les dificultats prosòdiques observades, fins i tot en els estudiants xinesos amb un domini avançat del castellà, ressalten la necessitat d'utilitzar paradigmes adequats d'entrenament prosòdic en la pràctica docent.

Paraules clau: Prosòdia; Castellà L2; Mandarí L1; Contorns nuclears; Segona llengua

Contornos nucleares de oraciones declarativas e interrogativas en español producidos por hablantes de mandarín con un dominio avanzado del español (Español)

Resumen: En este estudio se modelan dinámicamente los contornos nucleares de oraciones en español producidas por 16 estudiantes hablantes de mandarín y 9 hablantes nativos de español, utilizando un corpus de habla obtenido mediante una *discourse completion task*. Las oraciones investigadas incluyen declarativas (foco amplio, categórico y correctivo) e interrogativas absolutas (informativas, confirmatorias y preguntas con apéndice confirmativo). Nuestros resultados indican que los estudiantes chinos (a) no logran producir las configuraciones nucleares correctas para diferenciar las oraciones declarativas categóricas de las de foco correctivo; (b) pronuncian significativamente acentos tonales más altos en posición nuclear y acentos tonales más bajos en las sílabas átonas en comparación con los hablantes nativos; y (c) emiten tonos de frontera incorrectos debido a la posición del acento léxico en las palabras nucleares. Este estudio aporta evidencias empíricas a la teoría del aprendizaje prosódico de la L2 y remarca la importancia de los detalles fonéticos por encima de las similitudes/diferencias fonológicas entre hablantes de L1 y L2. Además, las dificultades prosódicas observadas, incluso en estudiantes chinos con un nivel avanzado de español, resaltan la necesidad de utilizar paradigmas adecuados de entrenamiento prosódico en la práctica docente.

Palabras clave: Prosodia; Español L2; Mandarín L1; Contornos nucleares; Segunda lengua

1. Introduction

In second language (L2) speech acquisition studies, much attention has been devoted to specific L2 speech sounds, with a rich output of empirical studies and theoretical accounts. However, L2 prosody has been neglected due to its dynamic nature and abstractness. Recent research has started filling this gap by providing empirical evidence and attempting to develop theoretical models on L2 prosody. Under this broad context, this study focuses on a less investigated language combination—L1 Mandarin speakers learning Spanish as an L2—to provide empirical evidence of L2 prosody in speech production. To analyze intonation, we used a non-linear statistical method—the Generalized Additive Mixed Model (GAMM)—to dynamically capture the fine-grained phonetic differences between L1 and L2 Spanish nuclear pitch contours.

1.1. Crosslinguistic transfer in L2 speech acquisition

The crosslinguistic transfer between L2 learners' first language (L1) and the target L2 is evidenced by research on L2 speech sounds. For instance, aspirating languages, such as Mandarin Chinese and English, contrast the stop consonants with [±aspiration] feature, while voicing languages like Spanish and Italian, show a [±voice] contrast (Lisker & Abramson, 1964). Accordingly, L1 speakers (L1ers) of aspirating languages produce [+voice] as [-aspiration] in voicing L2 (Feng & Busà, 2022; Li & Ye, 2022; Liu, 2016; Xi & Li, 2022, 2023), while voicing L1ers assimilate [+aspiration] to [-voice] in aspirating L2 (Flege & Eefting, 1987; Gorba & Cebrian, 2021; Li et al., 2021; Xi et al., 2020). Similarly, learners may assimilate an L2 vowel that does not exist in their L1 inventory to an L1 vowel. For instance, when learning French, Spanish speakers assimilate the French /y/ to Spanish /u/ due to the absence of /y/ in Spanish (Li, Baills, Baqué, et al., 2023; Racine & Detey, 2019; Turner, 2023). Likewise, Spanish-Catalan bilinguals may not distinguish the English /æ/ and /ʌ/ since these sounds are not part of Spanish or Catalan phonetic inventories (Cebrian, 2019; Xi et al., 2023; Xi, Li, et al., 2024). These findings align with the Speech Learning Model (SLM) and its revised version (SLM-r), which predict that learners establish an equivalence classification between L2 and L1 categories (Flege, 1995; Flege & Bohn, 2021).

Nevertheless, the theoretical frameworks of SLM or SLM-r, although widely supported by numerous empirical data, are explicitly proposed for the learning of individual sounds (Flege & Bohn, 2021, p. 3). The hypothesis within SLM/SLM-r might provide some insight into L2 prosodic learning, but its generalization requires further investigation. Mennen (2015) proposed a tentative framework to account for the learning of L2 prosody—the L2 Intonation Learning Theory (LILt). LILt, closely aligned with the theoretical accounts of SLM, posits that the similarities and dissimilarities between the learners' L1 and L2 prosodic inventories are the key predictor of successful L2 prosodic learning. For instance, Mandarin *yn*-questions are not necessarily marked by a rising boundary tone at the end of the sentence, especially when there is a question particle (Shao, 2014). This contrasts with Spanish, where *yn*-questions typically show a rising boundary tone (Estebas-Vilaplana & Prieto, 2010). Due to crosslinguistic influence, Mandarin L1ers tend to produce L2 Spanish *yn*-questions with a low-rising, or to a lesser degree, a non-rising intonation (Zhao, 2019, 2023; Zhao & Font-Rotchés, 2022).

Despite the systematic differences at a phonological level, other dimensions may also account for the L2 prosodic learning. First, the fine-grained phonetic details of phonological categories in prosody affect the outcome of L2 prosodic learning. For instance, although information-seeking *yn*-questions in both Spanish and Mandarin are realized with a rising boundary tone (LH%), Mandarin learners of Spanish tend to produce a higher pitch for LH% compared to native speakers (Shang & Elvira-García, 2022). Second, compatible with SLM, LILt considers that the age of acquisition is an important predictor of successful learning. That is, early exposure to an L2 typically results in a more target-like L2 prosody. Finally, in line with SLM, LILt assumes that as learning experience accumulates, L2 learners would approximate to the native prosodic norms. In short, LILt is largely inspired by SLM in terms of the production of L2 prosodic features. However, as the author pointed out, LILt had not been "tested directly as yet" (Mennen, 2015, p. 178). Therefore, more empirical evidence is needed to validate and refine this model. More importantly, from a crosslinguistic point of view, investigating language pairs with typologically different prosody will provide valuable evidence for the theoretical account of L2 prosodic learning. Particularly in the sense of crosslinguistic transfer, it is not sufficient to merely focus on the phonological aspects of prosody; the fine-grained phonetic details are also important. Therefore, in this study, we focus on how native speakers of Mandarin (a tone language) produce the prosody of Spanish (an intonation language), with a focus on fine-grained phonetic details using a dynamic analytic scheme. In the next section, we will review the literature on L2 Spanish prosodic learning by Mandarin L1ers.

1.2. L2 Spanish prosody acquisition by Mandarin-speaking learners

Spanish is an "intonation language", where pitch configurations are specified at the post-lexical level (Armstrong, 2015). By contrast, Mandarin is a "tone language" which configures pitch contours at the lexical level to differentiate word meanings (Zerbian et al., 2009). Moreover, Mandarin speakers use multiple prosodic features, like duration, to convey intonational nuances (Chen & Gussenhoven, 2008; Xi, Zhou, et al., 2024). Given these cross-linguistic differences in prosody, Mandarin L1ers tend to transfer Mandarin prosodic characteristics to L2 Spanish production. For instance, in running speech, Mandarin L1ers realize Spanish stressed syllables as Tone 2, which shows a rising pitch contour, while unstressed syllables are realized as Tone 4, which has a falling pitch contour (Chen, 2007). Recent studies showed differences between Mandarin L1ers and Spanish L1ers in how they use prosodic cues to mark Spanish lexical stress (Li & Xi, 2022, 2024). Specifically, Mandarin L1ers tend to enlarge pitch differences for marking the lexical stress contrasts in L2 Spanish, whereas Spanish L1ers rely more on durational differences. In short, even though Mandarin speakers have acquired the L2 Spanish lexical stress contrasts on phonological level, they show non-native phonetic details in making the phonological contrasts.

Nevertheless, only a couple of studies have investigated the phonetic realization of L2 Spanish intonation by Mandarin speakers. Cortés Moreno (2004) conducted an acoustic analysis to examine how Mandarin speakers transfer their L1 prosodic patterns to L2 Spanish. The author focused on four types of Spanish sentences: statements, *wh*-questions, *yn*-questions, and exclamations, and observed that the learners showed much higher pitch in declaratives than Spanish natives. In addition, the complexity of the intonation patterns in both Mandarin and Spanish seems to account for the learning difficulties. The most complex intonation, found in exclamatory sentences, is the most challenging, followed by less complex *yn*-questions, then *wh*-questions, and finally, statements (Cortés Moreno, 1997, 2005). Note that Cortés Moreno's series of work focused on phonological descriptions of the sentences produced by their informants. It remains unclear whether there are significant differences in fine-grained phonetic details between L2 learners and L1 speakers. Moreover, the four sentence types involved in the experimental design may show considerable variants in prosodic structures depending on pragmatic functions. For instance, Estebas-Vilaplana and Prieto (2010) identified four basic types of Spanish sentences, namely, statements, questions (*yn*-question and *wh*-question), imperatives, and vocatives. Statements and questions can be divided into neutral and biased depending on whether any kind of bias is included. Following this, statements can be broad-focused (neutral) or narrow-focused (biased), and *yn*-questions can serve for information-seeking (neutral) or confirmation-seeking (biased). The surface intonation structure varies depending on the specific pragmatic function, which may lead to extra challenges for L2 learners.

Recently, Shang and Elvira-García (2022) exclusively investigated the nuclear configurations of Spanish questions produced by Mandarin speakers, with various syntactic structures (*yn*-, *wh*-, tag, and disjunctive questions) and pragmatic functions (information-seeking vs. confirmation-seeking). They used the traditional analytic framework Sp_ToBI (Estebas-Vilaplana & Prieto, 2010) to phonologically categorize the nuclear pitch contours of the target sentences. They found that Mandarin speakers tend to produce the nuclear pitch accent in questions with a high (H*) or rising (L+H*) contour, instead of the expected low (L*) or falling (H+L*) contour. This suggests a transfer from L1 Mandarin, as none of the four Mandarin lexical tones is realized as a stable low tone. More importantly, Mandarin speakers often end Spanish interrogatives with a high boundary tone (H%) rather than a low boundary tone (L%), which may also result from crosslinguistic transfer. However, phonological labelling using the ToBI system relies on researchers' estimations, which inevitably introduces variances due to the researchers' degree of freedom. Recent developments in statistic methods, such as the Generalized Additive Mixed Model (GAMM) (Sóskuthy, 2021), allows for the dynamic analysis of pitch contour, considering changes in pitch values over time. This new method has yet to be widely introduced into L2 prosodic research practice. GAMM can compare the intonation curves produced by L2 learners to that of L1 speakers and statistically estimate the exact location where L2 speakers differ from L1 speakers. This method provides more robust and informative data than categorically labeling the prosodic features with abstract phonological representations.

Furthermore, another important factor that may account for L2 prosodic production is the interplay between lexical stress and intonation. For instance, as Mandarin speakers often map the stressed syllable to a rising tone (Chen, 2007), a stressed syllable might lead to a non-native high boundary tone or rising pitch contour. Considering this, Li and Xi (2023) used GAMM to analyze the nuclear pitch contours of Spanish echo questions produced by Mandarin-speaking learners. They examined three types of stress positions in the nuclear words: word-initial stress, word-medial stress, and word-final stress. The target sentences varied in syntactic structure (*yn*- and *wh*-question) and pragmatic function (information-seeking, clarification echo, and counter-expectation echo). Through the comparison between Mandarin speakers and Spanish natives, the GAMM analysis revealed that learners tend to replace the low boundary tone with a high boundary tone, aligning with Shang and Elvira-García's findings (2022). More importantly, the position of lexical stress played a role in the learning difficulties—the nuclear word with word-final stress was the most challenging for the learners.

1.3. The current study

Despite the fruitful findings from research on L2 Spanish prosodic acquisition, several research gaps have been identified. First, to advance theories of L2 prosody, more empirical data on fine-grained phonetic details are needed. Relying solely on phonological analysis overlooks crucial phonetic nuances (Sóskuthy, 2021; Wieling, 2018), especially since L2 speakers often present more variance than L1 speakers in the phonetic realizations of prosodic structures. Second, Spanish features diverse intonation contours for both syntactic and pragmatic purposes. Previous studies mainly considered the influence of syntactic structure on pitch contours (e.g., Cortés Moreno, 1997, 2004, 2005) or investigated the role of pragmatic function by focusing on specific sentence types (e.g., Shang & Elvira-García, 2022 on *yn*-questions), leaving many other sentence types unexplored. Therefore, a holistic description of how Mandarin-speaking learners produce L2 Spanish prosody should consider both syntactic and pragmatic elements. Third, only a few studies have examined the interplay between lexical stress and nuclear configuration (e.g., Li & Xi, 2023; Shang & Elvira-García, 2022), indicating a need for more research on the role of lexical stress position in L2 prosodic production.

To fill in the research gaps, the current study uses GAMM to dynamically analyze the prosodic patterns of Spanish sentences produced by Mandarin-speaking learners and Spanish native speakers. The data is part of a larger project examining the L2 prosody of Spanish sentences with varying syntactic structures and pragmatic functions. We mainly focus on two types of sentences: statements and *yn*-questions and only analyze the nuclear pitch contours of these sentences, as nuclear pitch accent and boundary tones are central to Spanish intonation (Prieto & Roseano, 2018). Additionally, we consider variations in the lexical stress of nuclear words to explore the stress-intonation interplay in L2 prosody.

1.3.1. The target sentence types

The sentences examined in this study—statements and *yn*-questions—are categorized into six subtypes, which vary in syntactic structure and pragmatic function, as listed below. The nuclear configurations for each sentence subtype are summarized based on Castilian Spanish (Estebas-Vilaplana & Prieto, 2010; Hualde & Prieto, 2015).

- Neutral statement is broad focus statement which carries all new information over the entire sentence. The most common nuclear configuration is L* L%.
- Biased statement introduces the speaker's belief or intention of highlighting a certain element, divided into two subtypes (Aguilar & Galicia, 2016):
    - Categorical focus statement expresses the speaker's strong belief towards the presupposition, which goes against what the interlocutor believes. The typical nuclear configuration is L* HL%.
    - Corrective focus statement is used to deny a specific element introduced in the topic and provides an alternative. The most common nuclear configuration is L+H* L%.
- Neutral *yn*-question, also called information-seeking *yn*-question, expresses that the speaker has no specific belief about whether the answer will be yes or no. The typical nuclear configuration is L* H%.
- Biased *yn*-question introduces the speaker's belief towards the truth condition of the presupposition encoded in the question. According to Estebas-Vilaplana and Prieto (2010), biased *yn*-questions are divided into echo questions and confirmation-seeking questions. In this study, we specifically focus on the confirmation-seeking questions with or without confirmatory tags.
    - Confirmation-seeking *yn*-question (confirmation question) encodes the speaker's belief towards a yes or no answer. There is no additional confirmatory tag attached to the sentence. Native speakers may realize it with either H+L* L% or L* H% in the nuclear position.
    - Tag question is pragmatically similar to confirmation-seeking question but includes a confirmative tag such as "*¿no*?", "*¿verdad*?", etc. In Shang and Elvira-García's (2022) corpus, 100% of the Spanish natives and Mandarin-speaking learners produced the confirmatory tag with L* H%, whereas the preceding word presented considerable variance. In our view, the word preceding the tag bears the pitch accent and boundary tone of an intermediate phrase which is also informative for L2 prosodic research. Therefore, in our study, we discarded the tags but treated the preceding word as the nuclear word for tag question.

The concrete scenarios for the six sentence subtypes, namely, neutral statement, categorical focus, corrective focus, neutral question, confirmation question, and tag question, are shown in Table 1 in the Materials subsection.

1.3.2. Research questions and hypotheses

Based on the literature review, we address the following research questions and hypotheses.

RQ1: Which subtype sentence(s) are the most challenging for Mandarin native speakers learning Spanish (henceforth, Chinese students)?

H1: Chinese students would face the greatest challenge in the biased sentences.

- For statements, the categorical focus would pose the greatest challenge due to the absence of the circumflex contour shape L* HL% in the Mandarin tone system (c.f. Shang & García-Elvira, 2022). Additionally, Chinese students may have difficulty distinguishing between the nuclear configurations of categorical focus (L* HL%) and corrective focus (L+H* L%), as both share a similar circumflex contour shape (Shang & García-Elvira, 2022).
- For *yn*-questions ("question" for short, thereafter), confirmation question would be the most challenging as Chinese students produce high boundary tones (H%) more frequently than Spanish natives (cf. Shang & Elvira-García, 2022). However, as previous research did not investigate the word preceding the confirmatory tags in tag question, no specific predictions will be made for this subtype.

RQ2: Do Chinese students realize the nuclear configurations of Spanish statements and *yn*-questions using different fine-grained phonetic details compared to Spanish natives? If so, what are the differences?

H2: Chinese students would replace the low pitch accent (L*) with high pitch accent (H*) (cf. Shang & Elvira-García, 2022) and produce the high pitch accent (H*) with an even higher pitch than Spanish natives (cf. Li & Xi, 2022).

RQ3: Which lexical stress position would affect the Chinese students' phonetic realization of Spanish nuclear configurations the most?

H3: The word-final stress position would be the most challenging when the boundary tone is supposed to be low (L%) (cf., Li & Xi, 2023). On the contrary, when the boundary tone is supposed to be high (H%), the word-initial or word-medial stress would be the least favorable.

2. Methods

2.1. Participants

We recruited 16 Chinese students (female = 12, aged 22-33 years) who are native speakers of Mandarin. These Chinese students had started learning Spanish since early adulthood (M = 19.3 years old, SD = 1.9). They had received formal Spanish instruction in China for 4.1 years (SD = 1.3) before moving to Spanish-speaking countries, where they had resided for 5.4 years (SD = 3.3). During their time in Spain, they had studied various subjects instructed in Spanish for 2.4 years (SD =2.2). In terms of their Spanish proficiency, all participants had passed the DELE test (Diplomas of Spanish as a Foreign Language) at either the B2 level (advanced, n = 5) or the C1 level (high advanced, n = 11). Notably, the five B2 participants had obtained the DELE B2 certificate before moving abroad, and their proficiency had increased beyond the B2 level through immersion and practical use in Spain, although none had formally taken the C1 test. Therefore, these Chinese students were classified as proficient late adult learners with substantial exposure to Spanish and long-term residence in the target L2 society.

As a baseline condition, we recruited 9 native speakers of peninsular Spanish (female = 6, aged 18-28 years). All Spanish natives grew up in Spain and reported speaking Spanish for their daily communication needs. Both groups of participants, the Chinese students and Spanish natives, signed written consent forms to allow the researchers to process their data.

2.2. Materials

Following Shang and García-Elvira (2022), we designed an adapted Discourse Completion Task (DCT)[1] with 54 trials to elicit Spanish utterances with different sentence structures and pragmatic functions. These include statements (broad, categorical, and corrective focus), *yn*-questions (information-seeking, confirmation-seeking, confirmation-seeking with tag, echo, request, and invitation), *wh*-questions (information-seeking, echo, request, and invitation), and imperatives (command and request). The current study only selected six target subtypes outlined in 1.3.1.

To explore the interaction between lexical stress and nuclear pitch contours, we designed three scenarios for each sentence subtype. Within each sentence subtype, each scenario elicited a sentence ending with one of the three target words, which varied in stress positions. These three target words were *vino* 'wine' (word-initial), *Marina* 'Marina' (word-medial), and *Milán* 'Milan' (word-final). This resulted in 18 scenarios. As the sentence-final target words bear the nuclear pitch contour, our analysis focused solely on the pitch contours of these three target words. Table 1 provides a comprehensive overview of all the materials along with their English translations.

Subtypes	Scenarios and target sentences	English translations
Neutral statement	(1) Lee la siguiente frase: *Elena bebe un vino*	(1) Read the following sentence: *Elena drinks a wine*
	(2) Lee la siguiente frase: *Noel va a Milán*	(2) Read the following sentence: *Noel goes to Milan*
	(3) Lee la siguiente frase: *Juanillo besó a Marina*	(3) Read the following sentence: *Juanillo kissed Marina*
Categorical focus	(1) Le has dicho a Julio que se añade *vino* en la carne a la Borgoña. Julio piensa que se añade salsa de soja. Dile, seguro/a, que no, que añadimos *vino*. *Que no, que añadimos vino*	(1) You've told Julio to add wine when cook beef Burgundy. Julio thinks that one should add soy sauce. Tell him with certainty that we add wine. *No, (I'm sure) we add wine*
	(2) Tu compañero te ha dicho que la nueva profesora española se llama Verónica, pero sabes que se llama Marina. Dile, seguro/a, que no, que se llama Marina. *Que no, que se llama Marina*	(2) Your friend told you the new Spanish teacher is called Veronica. But you know for sure that her name is Marina. Tell him with certainty that her name is Marina. *No, (I'm sure) her name is Marina*
	(3) Sabes que Eva y Juan irán a Milán, pero Mario piensa que irán a Lima. Dile, seguro/a, que no, que irán a Milán. *Que no, que irán a Milán*	(3) You know Eva and Juan will go to Milan. But Mario thinks they'll go to Lima. Tell him with certainty that they'll go to Milan. *No, (I'm sure) they'll go to Milan*
Corrective focus	(1) En una tienda ruidosa, Tú: Quiero dos botellas de *vino*. Dueño: ¿Cómo? ¿De agua? Tú: *No, de VINO*	(1) In a noisy store, You: I want two bottles of wine. Owner: Sorry? Of water? You: *No, of WINE*
	(2) En un aula ruidosa, Tú: Este libro es de Marina. Ema: ¿Cómo? ¿De María? Tú: *No, de MARINA*	(2) In a noisy classroom, You: This book is Marina's. Ema: Sorry? (It's) Maria's? Tú: *No, MARINA's*
	(3) En una fiesta ruidosa, Tú: La chica es de Milán. Ana: ¿Cómo? ¿De Lima? Tú: *No, de MILÁN*	(3) In a noisy party, You: The girl is from Milan. Ana: Sorry? From Lima? Tú: *No, from MILAN*

[1] To elicit target sentences that are acoustically analyzable, the target sentences for completing the discourse were given as a prompt, which distinguishes our design from a classic DCT used in prosodic studies (Vanrell Bosch et al., 2018).

Neutral question	(1) Estás en una tienda. Pregunta al dueño si tiene *vino*: *¿Tienes vino ?*	(1) You are in a store. Ask the owner if he has wine: *Do you have wine?*
	(2) Pregunta a tu compañero si conoce a Marina: *¿Conoces a Marina?*	(2) Ask your friend if he knows Marina: *Do you know Marina?*
	(3) Pregunta a tu compañero si le gusta Milán: *¿Te gusta Milán?*	(3) Ask your friend if he likes Milan: *Do you like Milan?*
Confirmation question	(1) Vas a una tienda que te parece que hay *vino*s en el estante. Pregunta al dueño para confirmar que tiene *vino*: *¿Tienes vino?*	(1) You go to a store that seems to have wine on the shelf. Ask the owner to confirm that he has wine: *You have wine?*
	(2) Acabas de llegar a Lima. Crees que estás en la calle donde está el hotel. Pregunta a alguien para confirmar que la calle es la Avenida de la Marina: *¿Es esta la Avenida de la Marina?*	(2) You just arrived in Lima. You think you are on the street where the hotel is located. Ask someone to confirm that the street is the Avenida de la Marina: *Is this the Avenida de la Marina?*
	(3) Encuentras a Stefano, quien es de Milán, en el aeropuerto de Madrid por casualidad. Crees que va a Milán. Le preguntas para confirmar: *¿Vas a Milán?*	(3) You meet Stefano, who is from Milan, at the Madrid airport by chance. You think he's going to Milan. You ask him to confirm: *You're going to Milan?*
Tag question	(1) Tu amigo te ha invitado a cenar a su casa, pero ya no tienes tiempo para comprar un *vino*. Esperas que tu amigo tiene *vino*. Se lo preguntas por teléfono para confirmar que tiene *vino*： *Tienes vino,¿ no?*	(1) Your friend has invited you to dinner at his place, but you don't have time to buy wine. You hope your friend has wine. You ask him on the phone to confirm that he has wine： *You have wine, don't you?*
	(2) Tienes que enviar un paquete a la casa de un amigo. Le preguntas por teléfono para confirmar que vive en la Avenida de la Marina: *Vives en la Avenida de la Marina, ¿verdad?*	(2) You need to send a package to a friend's home. You ask him by phone to confirm that he lives on Avenida de la Marina: *You live on the Avenida de la Marina, right?*
	(3) Recuerdas que tu amigo irá a Milán, pero no estás seguro/a de eso. Pregúntale para confirmar que va a Milán: *Vas a Milán, ¿verdad¿*	(3) You remember your friend is going to Milan, but you're not sure. Ask him to confirm that he is going to Milan: *You're going to Milan, right?*

Table 1. Experimental trials for the Discourse Completion Task. Target sentences are italicized.

2.3. Procedure

Participants were tested individually in a quiet room. The scenarios were organized in a PowerPoint presentation, with each slide containing one scenario followed by a prompt of the target sentence. All the trials were presented on a laptop in a randomized order. The experiment contained three phases: an introduction phase, an exercise phase, and an experimental phase.

In the introduction phase, participants were instructed on how they were expected to perform during the experiment. They were asked to carefully read the scenario descriptions and imagine the situations to the best of their ability. Participants were clearly instructed to read the non-target texts (i.e., the scenario descriptions) by heart. Then, using their imagination, they orally read the target sentences, which were in blue. If they happened to make any errors such as stumbled speech, hesitation, or mispronunciation, they were to correct the entire sentence immediately.

In the exercise phase, we provided the participants with three scenarios that were not part of the 54 experimental trials. Participants were allowed to practice as many times as necessary until they understood the experiment. They were also free to ask questions to the experimenters. The three exercise trials were also recorded but not analyzed. Once participants indicated that they fully understood the experimental procedure, they proceeded to the experimental phase.

In the experimental phase, participants were asked to carefully read by heart each scenario and orally read the target sentences in Spanish. Participants were self-paced. Upon finishing one trial, they slightly touched the touchpad of the laptop, and a cross sign would appear in the center of the screen. This was designed to adjust the participants' speed and to prevent unintentional skipping of PowerPoint slides due to a double click. The participants' speech outcomes were recorded to a Zoom H4n Pro recorder though a Shure SM35 headset condenser microphone.

2.4. Data coding and analysis

In total, we obtained 1350 utterances. For this study, we focused on the 6 target sentence subtypes, each containing one of the three target words with varied stress patterns (namely, *vino*, *Marina*, and *Milán*), resulting in 450 utterances (25 participants × 18 scenarios). To analyze the nuclear pitch contours, we manually segmented the target words at the syllable level using Praat (Boersma & Weenink, 2020). We then extracted 10 regularly spaced F0 samples from each syllable to generate a time-normalized pitch contour ("time" thereafter).

To estimate and compare the pitch contours of the target words across groups and pragmatic functions, we built six GAMMs using the *bam()* function from the *mgcv* package (Wood, 2017) in R for each of the two sentence structures (statement vs. question) across the three target words ("*vino*", "*Marina*", "*Milán*"). We closely followed the recommendations by Sóskuthy (2021) and Wieling (2018) to conduct the GAMM analysis. To minimize the intrinsic differences in pitch caused by individual factors such as gender, we transformed the raw pitch values into *z*-scores for each participant. This means that for each participant, the pitch values were centered at zero, which represents the mean pitch value of the participant him/herself.

For all six GAMMs, two sets of terms were included: the parametric terms and the smooth terms. The parametric terms included pragmatic, group, and Pragmatic × Group interaction. The baselines of pragmatic were the neutral "broad focus" for statement and the neutral "information seeking" for *yn*-question. The baseline of group was set to Spanish natives. Significant differences are interpreted in terms of mean pitch height, not the shape of the pitch contour.

The smooth terms included a smooth curve over time for the Pragmatic × Group interaction and a by participant random smooth over time to account for individual differences in pitch contour. The data distribution was specified as "scaled-t" to adjust the typical heavy-tailed data in pitch analyses (Chuang et al., 2021). For smooth terms, if the estimated degree of freedom (*edf*) is significantly different from 1, the estimated curve is said to be significantly different from a horizontal straight line. Given our interest in the differences in nuclear pitch contours as a function of speaker group and pragmatic meaning, we illustrated the estimated curves under comparison using the *plot_smooth()* function and plotted the differences between the compared curves using the *plot_diff()* function from the *itsadug* package (van Rij et al., 2022). Significant differences were labeled with red horizontal lines along the bottom of the plot.

3. Results

Figure 1 plots the pitch contours of each individual speaker, divided by pragmatic function and target word. There are considerable individual differences among participants in both the Spanish natives and Chinese students. More importantly, the visual inspection shows that the pitch values vary along the normalized time in a non-linear way, which is in favor of a non-linear model analysis. In the next sections, we report the GAMM results of each pragmatic function.

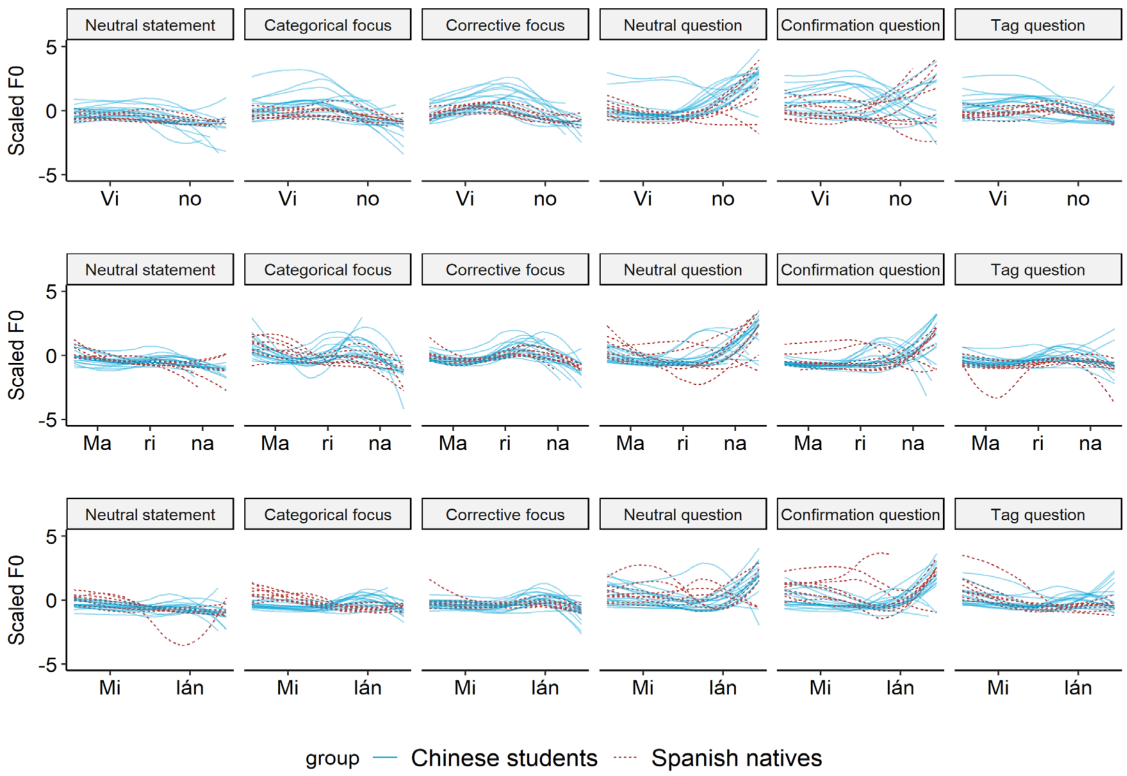

Figure 1. Pitch contour of the nuclear words produced by Chinese students and Spanish natives under six pragmatic functions. Each individual curve represents one participant.

3.1. Results of statements

3.1.1. Nuclear configurations of initial-stressed word in statements

Table 2 summarizes the model report of the initial-stressed "*vino*".

Parametric coefficients	**β**	***SE***	***t***	***p***	
(Intercept)	-0.64	0.12	-5.38	< .001	***
pragmatic [categorical]	0.27	0.04	7.02	< .001	***
pragmatic [corrective]	0.43	0.04	11.05	< .001	***
group [ch]	0.20	0.15	1.37	.172	
pragmatic [categorical] × group [ch]	0.26	0.05	5.49	< .001	***
pragmatic [corrective] × group [ch]	0.10	0.05	2.12	.034	*
Smooth terms	**edf**	**Ref. df**	***F***	***p***	
s(time) × interaction (neutral × sp)	1.00	1.00	6.88	.009	**
s(time) × interaction (categorical × sp)	2.98	3.70	4.81	.001	**
s(time) × interaction (corrective × sp)	4.60	5.64	12.82	< .001	***
s(time) × interaction (neutral × ch)	4.44	5.43	15.35	< .001	***
s(time) × interaction (categorical × ch)	6.34	7.46	35.11	< .001	***
s(time) × interaction (corrective × ch)	6.16	7.29	26.83	< .001	***
s(time, participant)	95.10	247.00	7.95	< .001	***
Adjusted R^2: 0.61, Deviance explained 0.54					

Table 2. GAMM summary of word-initial stressed "*vino*" in statements. Significant results are marked with asterisks: *p* < .05, *; *p* < .01, **; *p* < .001, ***.

We first interpret the parametric coefficients. The intercept is the mean pitch height of the target word "*vino*" produced by Spanish natives in neutral statement. The significant main effects of categorical focus and corrective focus indicate that Spanish natives produced significantly higher mean pitch in the categorical and corrective statements. There was no significant main effect of group, which means that Chinese students did not significantly differ from Spanish natives in mean pitch height in the neutral statement. However, the significant two-way interactions of Categorical focus × Ch and Corrective focus × Ch indicate that compared to the baseline condition, neutral statement by Spanish natives, Chinese students produced significantly higher mean pitch for "*vino*" in categorical and corrective foci.

The smooth terms all showed significant values. This means that the pitch contours of "*vino*" produced by both groups in all three conditions were significantly different from a horizontal straight line.

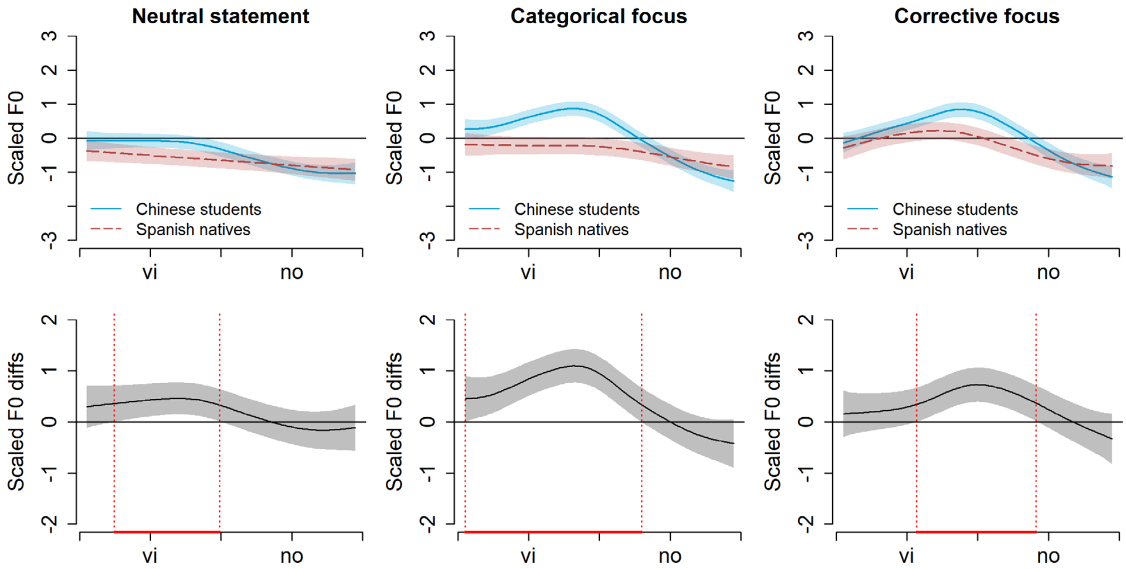

Figure 2. Estimated mean pitch contour and contour differences of word-initial stressed "*vino*" produced by Chinese students and Spanish natives in the nuclear position of statements under different focus conditions.

Finally, we visually inspect the post-hoc comparisons of pitch contours between Spanish natives and Chinese students. In Figure 2, the upper panels show the estimated pitch contours of "*vino*" produced by the two groups under the broad, categorical, and corrective foci, along with the 95% confidence intervals (CI) represented by shaded areas. The lower panels illustrate the significant contrasts between the two groups with red lines. Notably, in the neutral statement, Chinese students produced significantly higher pitch than Spanish natives on the stressed syllable "*vi*". In the categorical focus, Chinese students produced significantly higher pitch than Spanish natives throughout the stressed "*vi*" and parts of the unstressed "*no*". Lastly, in the corrective focus, both groups showed similar pitch patterns, although Chinese students showed significantly higher pitch during the latter half of the stressed "*vi*" and initial half of the unstressed "*no*".

3.1.2. Nuclear configurations of medial-stressed word in statements

Table 3 summarizes the model report of the medial-stressed "*Marina*".

In terms of the parametric coefficients, the significant main effects of categorical focus and corrective focus mean that Spanish natives produced significantly higher mean pitch for "*Marina*" in the categorical and corrective focus compared to the neutral statement. However, the pitch height of Chinese students did not significantly differ from that of Spanish natives in producing "*Marina*" in the neutral statement, as the group was not a significant predictor. Moreover, we only obtained one significant two-way interaction of Corrective focus × Ch, which suggests Chinese students produced significantly higher mean pitch for "*Marina*" in the corrective focus than did the Spanish natives in the neutral statement

.

Parametric coefficients	**β**	***SE***	***t***	***p***	
(Intercept)	-0.54	0.08	-6.77	< .001	***
pragmatic [categorical]	0.41	0.03	16.45	< .001	***
pragmatic [corrective]	0.35	0.03	14.16	< .001	***
group [ch]	-0.02	0.10	-0.24	.813	
pragmatic [categorical] × group [ch]	0.03	0.03	1.10	.270	
pragmatic [corrective] × group [ch]	0.11	0.03	3.67	< .001	***
Smooth terms	**edf**	**Ref. df**	***F***	***p***	
s(time) × interaction (neutral × sp)	1.00	1.00	41.98	< .001	***
s(time) × interaction (categorical × sp)	5.70	6.81	11.33	< .001	***
s(time) × interaction (corrective × sp)	6.17	7.28	22.32	< .001	***
s(time) × interaction (neutral × ch)	6.31	7.33	9.43	< .001	***
s(time) × interaction (categorical × ch)	8.00	8.66	21.84	< .001	***
s(time) × interaction (corrective × ch)	8.35	8.83	30.43	< .001	***
s(time, participant)	140.81	248.00	11.61	< .001	***
Adjusted R^2: 0.54, Deviance explained 0.52					

Table 3. GAMM summary of word-medial stressed "*Marina*" in statements. Significant results are marked with asterisks: $p < .05$, *; $p < .01$, **; $p < .001$, ***.

As for the smooth terms, all the smooth terms showed significant results, which means that the pitch contours of "*Marina*" in both groups across all conditions were significantly different from a horizontal straight line.

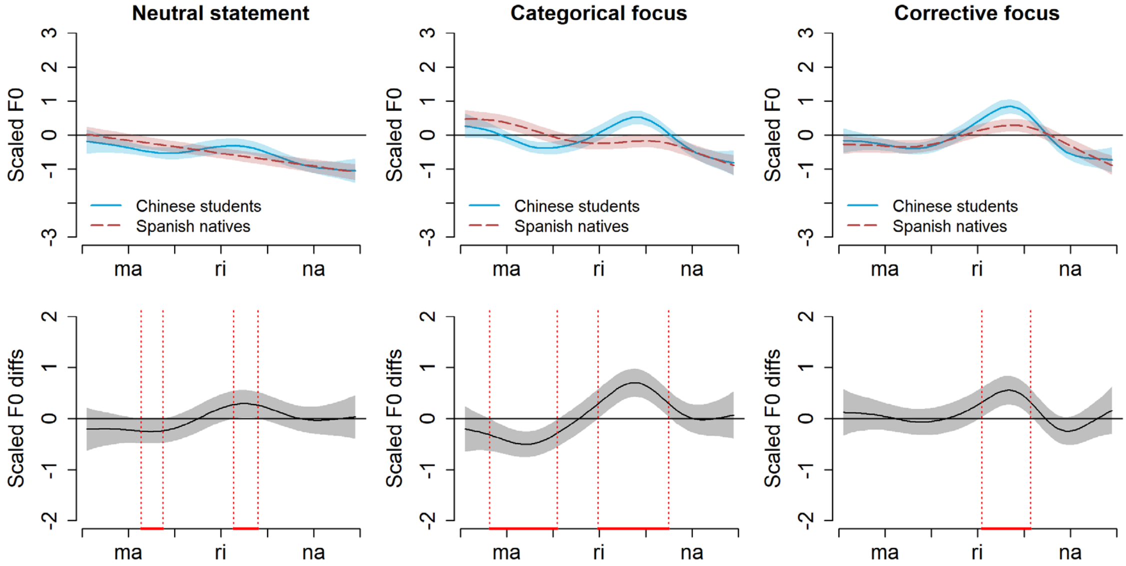

Figure 3. Estimated mean pitch contour and contour differences of word-medial stressed "*Marina*" produced by Chinese students and Spanish natives in the nuclear position of statements with different focus conditions.

The post-hoc comparisons are visualized in Figure 3. In the neutral statement, Chinese students produced quite similar pitch contours to those of Spanish natives. In the categorical focus, compared to Spanish natives, Chinese students produced significantly lower pitch on the unstressed "*ma*" but significantly higher pitch on the stressed "*ri*", resulting in a rising trend on the stressed syllable. Finally, in the corrective focus, both groups showed a high pitch on the stressed syllable "*ri*", but Chinese students produced even higher pitch than Spanish natives did on the second half of the stressed "*ri*".

3.1.3. Nuclear configurations of final-stressed word in statements

The GAMM results of the final-stressed "*Milán*" are reported in Table 4.

Parametric coefficients	**β**	***SE***	***t***	***p***	
(Intercept)	-0.46	0.06	-7.51	< .001	***
pragmatic [categorical]	0.29	0.02	13.34	< .001	***
pragmatic [corrective]	0.14	0.02	6.35	< .001	***
group [ch]	-0.16	0.08	-2.05	.040	*
pragmatic [categorical] × group [ch]	-0.03	0.03	-0.97	.333	
pragmatic [corrective] × group [ch]	0.19	0.03	7.01	< .001	***
Smooth terms	**edf**	**Ref. df**	***F***	***p***	
s(time) × interaction (neutral × sp)	2.39	2.84	27.16	< .001	***
s(time) × interaction (categorical × sp)	5.08	6.14	23.26	< .001	***
s(time) × interaction (corrective × sp)	5.22	6.29	4.82	< .001	***
s(time) × interaction (neutral × ch)	5.63	6.70	9.81	< .001	***
s(time) × interaction (categorical × ch)	7.93	8.65	25.85	< .001	***
s(time) × interaction (corrective × ch)	8.28	8.82	39.07	< .001	***
s(time, participant)	122.47	248.00	8.60	< .001	***
Adjusted R^2: 0.50, Deviance explained 0.53					

Table 4. GAMM summary of word-final stressed "*Milán*" in statements. Significant results are marked with asterisks: p < .05, *; p < .01, **; p < .001, ***.

First, categorical and corrective foci revealed significant main effects, suggesting that Spanish natives produced significantly higher mean pitch for "*Milán*" in the categorical and corrective foci than in the neutral statement. The significant main effect of group means that Chinese students produced "*Milán*" with significantly lower mean pitch than Spanish natives in the neutral statement. This finding contrasts with the results for "*vino*" and "*Marina*", where no significant group-level difference was found for the neutral statement. The significant two-way interaction of Corrective focus × Ch indicates that in the corrective focus, Chinese students produced significantly higher mean pitch for "*Milán*" than the Spanish natives did in the neutral statement.

Similar to "*vino*" and "*Marina*", the pitch contours of "*Milán*" produced by both Chinese students and Spanish natives were significantly different from a horizontal straight line across all focus conditions, as indicated by the significant smooth terms.

Finally, we interpret the post-hoc comparisons visualized in Figure 4. In the neutral statement, both Chinese students and Spanish natives produced similar downtrend pitch contours. There was a slightly lower pitch in the unstressed "mi" found in the Chinese students' speech. In the categorical focus, the Chinese students started with a significantly lower pitch than Spanish natives, yet they showed a high rise on the stressed syllable "*lán*". This sharp change resulted in a significantly higher pitch on the stressed syllable in the Chinese students' speech than in the Spanish natives' speech. In the corrective focus, the pitch contour patterns of both groups were similar to that of the categorical focus. However, significant contrasts were only observed on the stressed "*lán*", with Chinese students showing significantly higher pitch than Spanish natives.

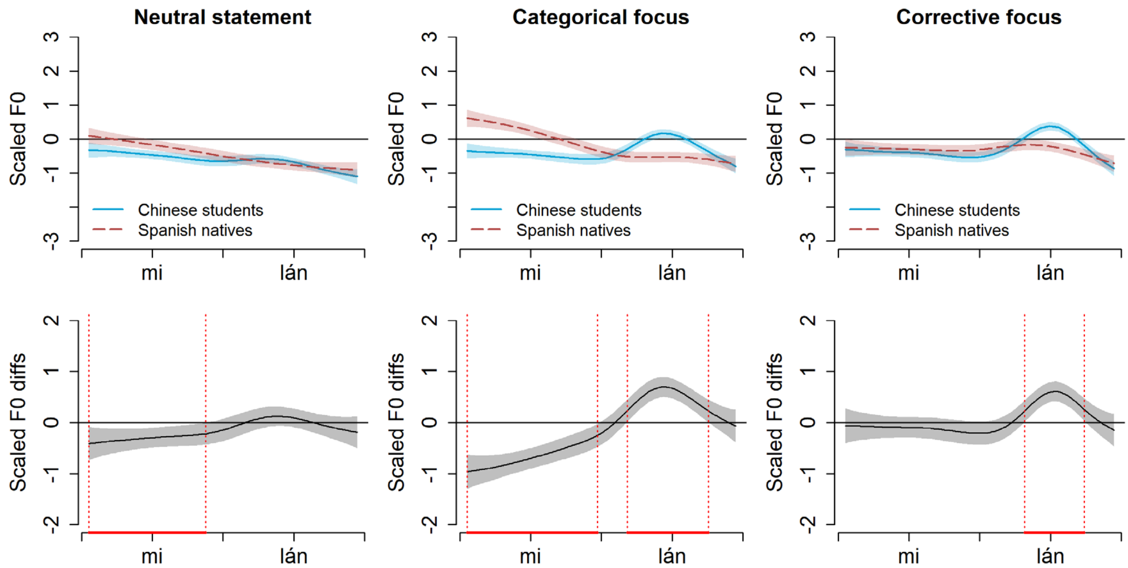

Figure 4. Estimated mean pitch contour and contour differences of word-final stressed "*Milán*" produced by Chinese students and Spanish natives in the nuclear position of statements with different focus conditions.

3.1.4. Interim summary of the results of statements

In this section, we briefly summarize the findings on the nuclear pitch contours of Spanish statements with different foci produced by Chinese students and Spanish natives.

First, the most challenging focus type for Chinese students is the categorical focus, as categorical focus triggered the largest group-level differences in pitch contours regardless of the nuclear word's stress position. By comparing Figures 2–4, it can be noted that Chinese students produced similar pitch contours for categorical and corrective foci, unlike Spanish natives who used distinct nuclear pitch contours. Interestingly, the corrective focus was not quite challenging for Chinese students to produce. They heightened the pitch of the stressed syllable to mark this focus type. The main difference came from Chinese students showing larger pitch range than Spanish natives rather than using a different contour shape.

Second, Chinese students produced stressed syllables with significantly higher pitch than Spanish natives in almost all cases. The only exception was "*Milán*" under the neutral statement, where no significant contrast was found. By contrast, Chinese students tended to produce unstressed syllables with significantly lower pitch than Spanish natives in many cases. This pattern was observed in the broad and categorical foci, where the syllable preceding the stressed syllables in "*Marina*" and "*Milán*" were lower in pitch in the Chinese students' speech compared to that of Spanish natives. This explains why Chinese students showed significant differences in the easiest neutral statement, as they realized stressed syllables differently from Spanish natives.

Third, the syllable-final stress position (*Milán*) was the least favorable for producing the correct boundary tone of categorical focus. Spanish natives produced categorical focus with a low boundary tone, whereas Chinese students produced it with a falling boundary tone, which was affected by the final stress.

3.2. Results of *yn*-questions

3.2.1. Nuclear configurations of initial-stressed word in yn-questions

Table 5 presents the GAMM report for the initial-stressed "*vino*" in *yn*-questions.

Parametric coefficients	**β**	***SE***	***t***	***p***	
(Intercept)	0.01	0.17	0.06	.956	
pragmatic [confirmation]	0.25	0.06	4.27	< .001	***
pragmatic [tag]	-0.19	0.06	-3.32	.001	**
group [ch]	0.69	0.22	3.18	.002	**
pragmatic [confirmation] × group [ch]	-0.24	0.07	-3.24	.001	**
pragmatic [tag] × group [ch]	-0.17	0.07	-2.29	.022	*
Smooth terms	**edf**	**Ref. df**	***F***	***p***	
s(time) × interaction (neutral × sp)	3.97	4.88	7.52	< .001	***
s(time) × interaction (confirmation × sp)	4.02	4.94	6.81	< .001	***
s(time) × interaction (tag × sp)	4.42	5.41	5.54	< .001	***
s(time) × interaction (neutral × ch)	4.46	5.49	26.45	< .001	***
s(time) × interaction (confirmation × ch)	1.00	1.00	0.00	.980	
s(time) × interaction (tag × ch)	4.55	5.59	13.56	< .001	***
s(time, participant)	92.66	246.00	9.33	< .001	***
Adjusted R^2: 0.386, Deviance explained 0.452					

Table 5. GAMM summary of word-initial stressed "*vino*" in *yn*-questions. Significant results are marked with asterisks: $p < .05$, *; $p < .01$, **; $p < .001$, ***.

We first focus on the parametric coefficients. The significant effects of confirmation question and tag question indicate that compared to the neutral question, Spanish natives produced the nuclear word "*vino*" with significantly higher mean pitch in the confirmation question, but significantly lower pitch in the tag question. The significantly lower pitch in the tag question may indicate that the interrogativity was conveyed by the question tags "*no*" or "*verdad*", possibly causing the *yn*-question to perform more like a statement. The significant main effect of group indicates that Chinese students produced "*vino*" in the neutral question with significantly higher pitch than Spanish natives. Furthermore, the significant two-way interactions of Confirmation question × Ch and Tag question × Ch indicate that compared to Spanish natives in the neutral question, Chinese students produced "*vino*" with significantly lower mean pitch in the nuclear position of the confirmation and the tag questions.

As for the smooth terms, only the pitch contour of "*vino*" produced by Chinese students in the confirmation question was not significantly different from a straight line. The rest of the conditions showed significant results.

Post-hoc comparisons are plotted in Figure 5. In the neutral questions, Spanish natives and Chinese students showed similar pitch patterns, but Chinese students had a significantly higher pitch rise at the second syllable "*no*" which resulted in a higher boundary tone. In the confirmation question, Chinese students showed a significantly higher pitch on the stressed syllable "*vi*", but their boundary tone was not rising, which resulted in a significantly lower boundary tone compared to Spanish natives. In the tag question, again, Chinese students showed a significantly higher pitch on the stressed syllable "*vi*". However, unlike the confirmation question, both Chinese students and Spanish natives showed low boundary tone at the end of "*vino*" in the tag question, with no significant contrast, resulting in a statement-like pattern.

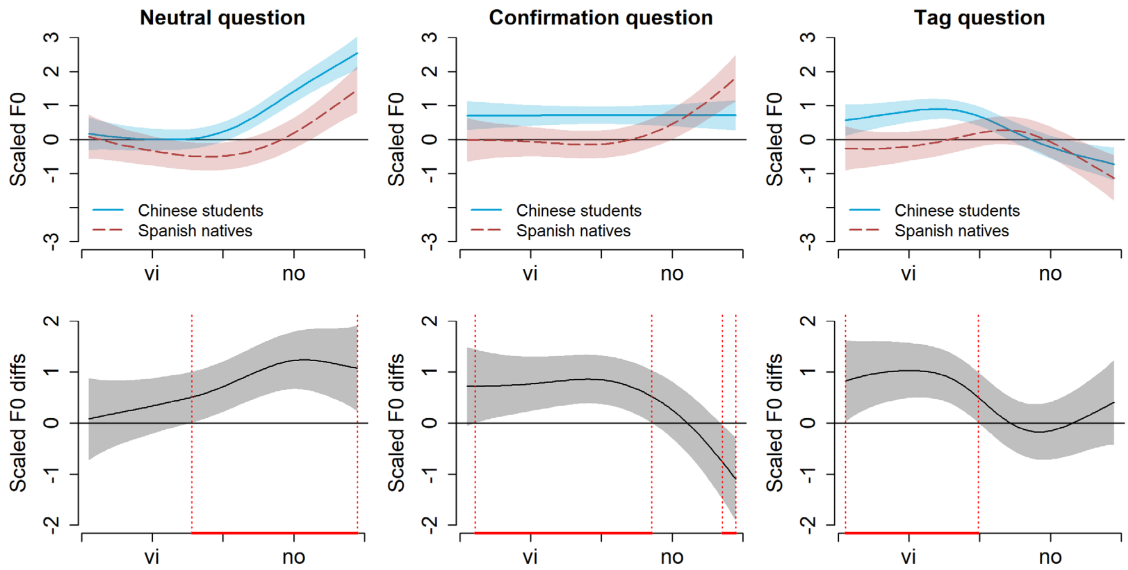

Figure 5. Estimated mean pitch contour and contour differences of word-initial stressed "*vino*" produced by Chinese students and Spanish natives in the nuclear position of *yn*-questions with different pragmatic functions.

3.2.2. Nuclear configurations of medial-stressed word in yn-questions

Table 6 shows the GAMM results of the medial-stressed "*Marina*" in *yn*-questions.

Parametric coefficients	**β**	***SE***	***t***	***p***	
(Intercept)	0.06	0.08	0.71	.479	
pragmatic [confirmation]	-0.34	0.03	-12.70	< .001	***
pragmatic [tag]	-0.56	0.03	-21.42	< .001	***
group [ch]	0.09	0.10	0.87	.383	
pragmatic [confirmation] × group [ch]	0.04	0.03	1.26	.210	
pragmatic [tag] × group [ch]	0.13	0.03	4.08	< .001	***
Smooth terms	**edf**	**Ref. df**	***F***	***p***	
s(time) × interaction (neutral × sp)	6.77	7.75	8.90	< .001	***
s(time) × interaction (confirmation × sp)	6.76	7.74	10.28	< .001	***
s(time) × interaction (tag × sp)	4.55	5.39	2.13	.055	
s(time) × interaction (neutral × ch)	5.92	6.90	16.70	< .001	***
s(time) × interaction (confirmation × ch)	6.40	7.39	14.89	< .001	***
s(time) × interaction (tag × ch)	6.66	7.62	7.56	< .001	***
s(time, participant)	161.00	246.00	27.67	< .001	***
Adjusted R^2: 0.52, Deviance explained 0.49					

Table 6. GAMM summary of word-medial stressed "*Marina*" in *yn*-questions. Significant results are marked with asterisks: *p* < .05, *; *p* < .01, **; *p* < .001, ***.

First, we interpret the parametric coefficients. The significant main effects of confirmation question and tag question indicate that Spanish natives produced "*Marina*" with significantly lower mean pitch in the confirmation and tag questions compared to the neutral question. However, there was no significant differences between Chinese students and Spanish natives in producing "*Marina*" for the neutral question, as group was not a significant factor. Moreover, the significant two-way interaction of Tag question × Ch suggests that Chinese students produced

"*Marina*" with significantly higher pitch in the tag question than Spanish natives did in the neutral question.

The smooth terms revealed that Spanish natives produced "*Marina*" in the tag question with a wiggling pitch contour, although it was not significantly different from a horizontal straight line due to its large standard error (p = .055). In all other conditions, "*Marina*" was produced with wiggling pitch contours.

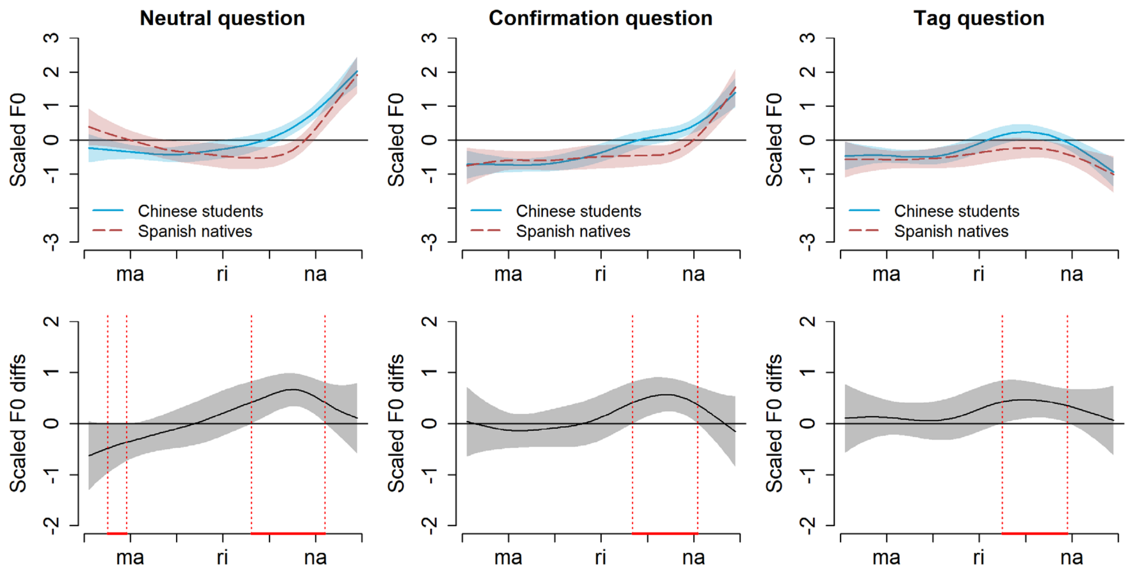

Figure 6. Estimated mean pitch contour and contour differences of word-medial stressed "*Marina*" produced by Chinese students and Spanish natives in the nuclear position of *yn*-questions with different pragmatic functions.

The post-hoc comparisons are presented in Figure 6. In the neutral question, both Chinese students and Spanish natives showed quite similar pitch contours for "*Marina*". Despite some subtle significant differences, both groups realized a rising pitch contour for the neutral question. Similar findings were observed for the confirmation question, where both groups produced "*Marina*" with an upward pitch contour. By contrast, "*Marina*" in the tag question was realized as a falling pitch contour by both groups. Again, the significant group-level contrasts were minor.

3.2.3. Nuclear configurations of final-stressed word in yn-questions

Table 7 summarizes the GAMM results of the final-stressed "*Milán*" in *yn*-questions.

In terms of the parametric coefficients, the significant main effect of tag question means that Spanish natives produced "*Milán*" with significantly lower pitch height in the tag questions compared to the neutral question. No group difference was found for the pitch height of "*Milán*" in the neutral question. Then, the significant two-way interactions of Confirmation question × Ch and Tag question × Ch indicate that Chinese students produced "*Milán*" with significantly lower mean pitch in the confirmation question but significantly higher mean pitch in the tag question, both compared to how Spanish natives produced "*Milán*" in the neutral question. Regarding the smooth terms, the pitch contours of "*Milán*" were significantly different from a horizontal straight line in all the conditions produced by both Chinese students and Spanish natives.

Parametric coefficients	**β**	***SE***	***t***	***p***	
(Intercept)	0.40	0.09	4.24	< .001	***
pragmatic [confirmation]	0.08	0.04	1.76	.079	
pragmatic [tag]	-0.58	0.04	-13.79	< .001	***
group [ch]	-0.13	0.12	-1.11	.266	
pragmatic [confirmation] × group [ch]	-0.30	0.05	-5.68	< .001	***
pragmatic [tag] × group [ch]	0.26	0.05	5.05	< .001	***
Smooth terms	**edf**	**Ref. df**	***F***	***p***	
s(time) × interaction (neutral × sp)	6.50	7.62	11.55	< .001	***
s(time) × interaction (confirmation × sp)	7.00	8.05	30.01	< .001	***
s(time) × interaction (tag × sp)	3.64	4.46	4.96	.001	**
s(time) × interaction (neutral × ch)	6.45	7.57	28.06	< .001	***
s(time) × interaction (confirmation × ch)	6.81	7.89	25.81	< .001	***
s(time) × interaction (tag × ch)	5.42	6.53	6.17	< .001	***
s(time, participant)	91.83	248.00	7.03	< .001	***
Adjusted R^2: 0.46, Deviance explained 0.48					

Table 7. GAMM summary of word-final stressed "*Milán*" in *yn*-questions. Significant results are marked with asterisks: $p < .05$, *; $p < .01$, **; $p < .001$, ***.

Finally, we interpret the post-hoc comparisons illustrated in Figure 7. Both groups of participants produced similar pitch contours in all three conditions. The significant differences were limited to parts of the syllables. Concretely, in the neutral question, compared to Spanish natives, Chinese students showed lower pitch in the unstressed "*mi*" but higher pitch in a very small proportion of the stressed "*lán*". However, in the confirmation question, Chinese students only showed significantly lower pitch in the unstressed "*mi*" than the Spanish natives. Lastly, in the tag question, Chinese students showed significantly higher pitch on the stressed "*lán*" than Spanish natives.

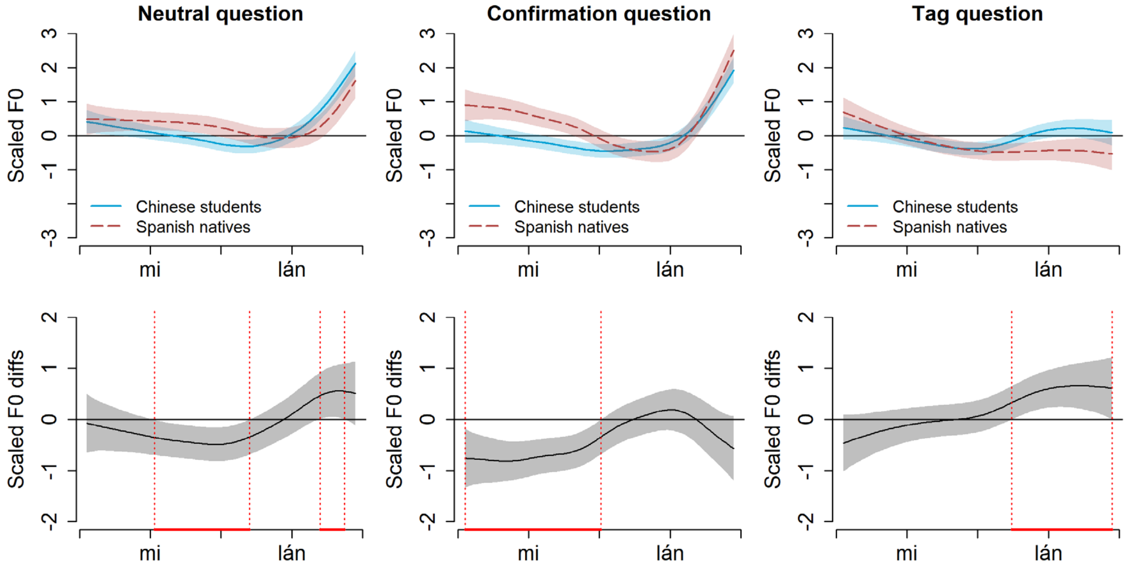

Figure 7. Estimated mean pitch contour and contour differences of word-final stressed "*Milán*" produced by Chinese students and Spanish natives in the nuclear position of *yn*-questions with different pragmatic conditions.

3.2.4. Interim summary of the results of yn-questions

First, both Chinese students and Spanish natives tended to use a rising boundary tone at the end of the neutral and confirmation questions, regardless of the stress position. However, there was one exception: Chinese students produced a low boundary tone for "*vino*" in the confirmation question, which contrasted with Spanish natives showing a rising boundary tone. As for the tag question, both groups consistently produced a low boundary tone for all three target words, suggesting that Chinese students had already acquired the correct boundary tone for tag questions.

Second, consistent to statements, when producing *yn*-questions, Chinese students showed significantly higher pitch in stressed syllables than Spanish natives with two exceptions: "*vino*" in the neutral question and "*Milán*" in the confirmation question. In terms of unstressed syllables, again, some syllables preceding stressed syllables were produced with significantly lower pitch by the Chinese students than by the Spanish natives. This was observed in "*Marina*" in the neutral question and "*Milán*" in the neutral and confirmation questions.

Third, the word-initial position (*vino*) did not seem to favour the production of a rising boundary tone in the confirmation question. Again, this may be affected by the unstressed syllable "*no*" being realized with a low tone.

4. Discussion

This study assessed the production of Spanish nuclear configurations in both statement and *yn*-questions, comparing the performance of Chinese students with that of Spanish natives using an adapted discourse completion task. The target nuclear words varied in stress positions (initial, medial, and final stressed), and the target sentences varied in pragmatic functions (neutral statement, categorical focus, corrective focus, neutral question, confirmation question, and tag question). The results unfolded several key findings. First, Chinese students had difficulties in prosodically distinguishing categorical focus from corrective focus, as well as in producing confirmation question. In other words, biased statements or *yn*-questions are more challenging than neutral ones. Second, compared to Spanish natives, Chinese students tended to produce stressed syllables with higher pitch and unstressed syllables with lower pitch, which was potentially affected by the heavy functional load of pitch specifications at the word level in their L1, Mandarin. Third, the positions of lexical stress had important impact on the phonetic realizations of Spanish boundary tones by Chinese students. Specifically, word-initial stress was not helpful for a high or rising boundary tone. The effects of lexical stress can be explained by the fact that Chinese students tend to associate a rising or high tone with stressed syllables and a falling or low tone with unstressed syllables (Chen, 2007). In what follows, we will organize our discussion to address each RQ, followed by an exploration of the theoretical and practical implications, as well as the limitations and directions for future research.

4.1. The phonetic realization of Spanish nuclear configurations in L2 and L1 speech

RQ1 aimed to identify the most challenging sentence subtypes for Chinese students, with H1 predicting that the biased conditions would pose more challenge than neutral conditions. Concretely, we predicted that the categorical focus and the confirmation question would be the most challenging. Our data validated H1 regarding statements, revealing that Chinese students produced the categorical focus similarly to the corrective focus. We also predicted that the circumflex contour shape would be challenging for Chinese students due to its absence in Mandarin lexical tones. However, our data showed that Chinese students could correctly produce this contour shape. Surprisingly, in our corpus, Spanish natives produced the categorical focus with a nuclear contour similar to the neutral statement, contrary to previous findings. Our discourse completion task was a reading task, while the previous studies (Estebas-Vilaplana &

Prieto, 2010; Hualde & Prieto, 2015; Prieto & Roseano, 2018) collected spontaneous speech data. This methodological difference may have influenced Spanish natives to orally produce the categorical focus in a neutral statement manner. As for *yn*-questions, we found minor differences in pitch contours between Chinese students and Spanish natives, except for "*vino*" in the confirmation question. This finding partially supports H1 predicting that biased conditions would be challenging for Chinese students. Finally, the confirmation question with tag posed no difficulty for Chinese students. This aligns with the findings of Shang and Elvira-García (2022), even though our focus was on the words preceding the confirmatory tags, while their study focused on the confirmatory tags themselves.

RQ2 asked whether Chinese students realize the nuclear configurations of Spanish statements and *yn*-questions using different fine-grained phonetic details compared to Spanish natives, and if so, what those differences are. Our data validated H2 which predicted that Chinese students would use pitch differences to mark lexical stress to a larger extent than Spanish natives. Specifically, we found that in most cases, Chinese students produced higher pitch on stressed syllables and lower pitch on unstressed syllables compared to Spanish natives. This finding replicated previous research (Li & Xi, 2022) and extended the pattern from isolated words to nuclear words. Moreover, the tendency to use high pitch for stressed syllables led to incorrect pitch accent configurations. Specifically, if a pitch accent should be low (L*) or not sufficiently high, it may be realized as high (H*) in L2 speech. This is well in line with Shang and Elvira-García's (2022) findings and expanded their conclusion from questions to statements. It is plausible that a similar pattern may be observed in other L2 Spanish sentence types, though empirical data are needed to confirm this hypothesis.

Finally, RQ3 investigated which lexical stress position would affect the Chinese students' phonetic realizations of Spanish nuclear configurations the most. We hypothesized that (H3) word-final stress would pose a challenge for realizing the low/falling boundary tone (HL% or L%) while word-initial or -medial stress would affect the realization of the high boundary tone (H%). Our data partly confirmed H3 that at least the word-initial stressed "*vino*" was realized as L% by Chinese students in confirmation *yn*-questions, contrary to LH% in L1 speech. On the other hand, word-final stress did not affect the realization of L% or HL%. This can be explained by the fact that statements are typically produced with L% in both Chinese and Spanish, regardless of being biased or neutral. Therefore, it is not challenging for Chinese students to correctly produce the expected boundary tone in L2 Spanish. By contrast, previous studies observed that Spanish natives showed various boundary tones in confirmation *yn*-questions (H+L* L% or L* H%, Estebas-Vilaplana & Prieto, 2010). Therefore, the variation may have posed challenges for L2 learners.

4.2. Theoretical and practical implications

First, we showed that the difficulties in learning L2 prosody on phonological level are largely driven by pragmatic functionality. Our findings revealed that pragmatically biased conditions are more challenging than neutral conditions. When learners struggle to grasp a certain pragmatic function, its corresponding prosodic structure in L2 speech production tends to be less on target. This is evident in our study where Chinese students could not prosodically mark Spanish categorical and corrective focus. In practice, L2 teachers may want to design engaging and interactive training materials that embed the training targets within meaningful context, which has been intensively used in recent research on L2 prosodic training (Jiang & Chun, 2024; Li et al., 2023, 2024; see also Levis, 2024, for a comprehensive summary).

Second, our study revealed a clear crosslinguistic transfer from L1 to L2. Concretely, Chinese students tended to realize nuclear pitch accent with higher pitch than Spanish natives. This finding is crucial as it largely accounts for the nonnative prosodic patterns found in Mandarin speakers' L2 speech. At least in our corpus, most significant differences between Chinese students and Spanish natives were pitch height differences, with Mandarin speakers producing unexpected high pitch for nuclear pitch accents or overly low pitch for unstressed syllables. This finding adds clear evidence to the LILt, which posits that “L1 influences are not solely restricted to the level of phonological contrasts” (Mennen, 2015, p. 179) and highlights the importance of L2 learners’ phonetic realizations of prosodic structures. This finding also provides insights into L2 prosodic training paradigms. Particularly, Mandarin native speakers might benefit from training on their pitch height to achieve finer-grained phonetic details in L2 prosodic production. Hence, L2 teachers are encouraged to incorporate pitch adjustment strategies into their teaching syllabi.

Finally, the interplay between lexical stress and intonation was validated by our data as well. In L2 prosodic research, it is therefore important to consider the interaction between lexical-level prosody and intonation-level prosody in experimental design. In L2 teaching practice, we propose a three-step approach for training Spanish nuclear configurations: (a) selecting target words with facilitative lexical stress positions for the training target (e.g., word-final stress for a rising boundary tone); (b) once learners could correctly produce the boundary tones under the facilitative lexical stress positions, introducing nuclear words with non-facilitative lexical stress patterns (e.g., word-initial stress for a rising boundary tone); and (c) providing explicit corrections for any mistakes, especially those caused by non-facilitative lexical stress positions. Nevertheless, due to space limitations, we are not able to provide detailed teaching recommendations. A future study might want to test the above-mentioned teaching proposals.

5. Conclusion

To conclude, this study is one of the first to focus on the fine-grained phonetic details of L2 Spanish (an intonation language) prosody produced by L1 Mandarin (a tone language) speakers using a dynamic analytic scheme. We showed that the learning difficulties at both phonological and phonetic levels are derived from the prosodic characteristics of learners’ L1 and the interplay between lexical stress and intonation. Our results hold significant implications for the development of L2 prosodic theories and L2 prosodic teaching practice. Notably, despite our participants’ extensive L2 experience and advanced proficiency in the target language, they still showed nonnative prosodic patterns. This calls for the needs of proper prosodic training paradigms in L2 teaching research.

Despite the positive findings, the current study suffers from several limitations. First, we exclusively focused on L2 Spanish learners with long-term study abroad experience in Spanish-speaking countries. Future studies should consider including learners with varying proficiency levels, as L2 experience may play a role in prosodic learning. Second, the current study was a one-time observation. However, with proper training paradigms, learners could improve their prosodic patterns over time. Moreover, implementing longitudinal observations will also provide insightful data for the development of L2 prosodic learning. Lastly, a crosslinguistic design involving intonation L1ers learning tone L2 and tone L1ers learning intonation L2 would yield more robust results to advance the crosslinguistic transfer on prosodic level.

Acknowledgements

This work is supported by the Research Council of Norway through its Centers of Excellence funding scheme [223265]. The authors sincerely thank Ms. Paula G. Sánchez-Ramón (Universitat Pompeu Fabra) for proofreading the Catalan and Spanish abstracts.

References

Aguilar Ruiz, M. C., & Galicia Isasmendi, I. (2016). La relación prosódico-gestual en los enunciados de foco contrastivo y declarativos categóricos en el español de Puebla, México. *Estudios de Lingüística Aplicada*, *63*, 37-58.

Armstrong, M. E. (2015). Accounting for intonational form and function in Puerto Rican Spanish polar questions. *Probus*, *2015*, 1–40. https://doi.org/10.1515/probus-2014-0016

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

Cebrian, J. (2019). Perceptual assimilation of British English vowels to Spanish monophthongs and diphthongs. *The Journal of the Acoustical Society of America*, *145*(1), EL52–EL58. https://doi.org/10.1121/1.5087645

Chen, Y. (2007). From tone to accent: The tonal transfer strategy for Chinese L2 learners of Spanish. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Science* (pp. 1645–1648). Saarbrücken Univ. des Saarlandes.

Chen, Y., & Gussenhoven, C. (2008). Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics*, *36*, 724–746. https://doi.org/10.1016/j.wocn.2008.06.003

Chuang, Y.-Y., Fon, J., & Baayen, R. H. (2021). Analyzing phonetic data with generalized additive mixed models. In M. Ball (Ed.), *Manual of Clinical Phonetics*. Routledge.

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

Cortés Moreno, M. (2004). Análisis acústico de la producción de la entonación española por parte de sinohablantes. *Estudios de Fonética Experimental*, 13, 79–110.

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

Estebas-Vilaplana, E., & Prieto, P. (2010). Castilian Spanish intonation. In P. Prieto & P. Roseano (Eds.), *Transcription of Intonation of the Spanish Language* (pp. 17–48). LINCOM publishers.

Feng, Q., & Busà, M. G. (2022). Acquiring Italian stop consonants: A challenge for Mandarin Chinese-speaking learners. *Second Language Research*. 29 (3) https://doi.org/10.1177/02676583221079147

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 233–277). York Press.

Flege, J. E., & Bohn, O.-S. (2021). The Revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second Language Speech Learning: Theoretical and Empirical Progress* (pp. 3–83). Cambridge Univesity Press.

Flege, J. E., & Eefting, W. (1987). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, *15*(1), 67–83. https://doi.org/10.1016/S0095-4470(19)30538-8

Gorba, C., & Cebrian, J. (2021). The role of L2 experience in L1 and L2 perception and production of voiceless stops by English learners of Spanish. *Journal of Phonetics*, *88*, 101094. https://doi.org/10.1016/j.wocn.2021.101094

Hualde, J. I., & Prieto, P. (2015). Intonational variation in Spanish: European and American varieties. In S. Frota & P. Prieto (Eds.), *Intonation in Romance* (pp. 450–391). Oxford University Press.

Jiang, Y., & Chun, D. M. (2024). The pedagogy of training Chinese students' intonation online. *Journal of Second Language Pronunciation*, 9 (3), 332-339.https://doi.org/10.1075/jslp.23045.jia

Levis, J. M. (2024). Teaching prosody in research studies: How researchers make decisions about pedagogy in research studies. *Journal of Second Language Pronunciation*, 9(3), 285-288. https://doi.org/10.1075/jslp.23047.lev

Li, P., Baills, F., Alazard-Guiu, C., Baqué, L., & Prieto, P. (2023). A pedagogical note on teaching L2 prosody and speech sounds using hand gestures. *Journal of Second Language Pronunciation*, *9*(3). https://doi.org/10.1075/jslp.23043.li

Li, P., Baills, F., Baqué, L., & Prieto, P. (2023). The effectiveness of embodied prosodic training in L2 accentedness and vowel accuracy. *Second Language Research*, *39*(4), 1077–1105. https://doi.org/10.1177/02676583221124075

Li, P., & Xi, X. (2022). Spanish lexical stress produced by proficient Mandarin learners of Spanish. *Proceedings of the 4th International Symposium on Applied Phonetics* (pp.41–45). https://doi.org/10.21437/ISAPh.2022-8

Li, P., & Xi, X. (2023). Nuclear contours of Spanish echo questions produced by proficient Chinese learners of Spanish: A dynamic analysis. In R. Skarnitzl & J. Volín (Eds.), *Proceedings of the 20th International Congress of Phonetic Sciences* (pp. 1201–1205). Guarant International.

Li, P., & Xi, X. (2024). The perception of Spanish lexical stress by proficient Mandarin learners of Spanish. *Proceedings of the 12th International Conference on Speech Prosody* (354-358). https://doi.org/10.21437/SpeechProsody.2024-72

Li, P., Xi, X., Baills, F., & Prieto, P. (2021). Training non-native aspirated plosives with hand gestures: Learners' gesture performance matters. *Language, Cognition and Neuroscience*, *36*(10), 1313–1328. https://doi.org/10.1080/23273798.2021.1937663

Li, P., & Ye, C. (2022). Voice onset time of Japanese stops produced by Shanghainese-Mandarin learners with various Japanese proficiency levels. In E. Babatsouli (Ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2022* (pp. 152–157). University of Louisiana at Lafayette.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *WORD*, *20*(3), 384–422. https://doi.org/10.1080/00437956.1964.11659830

Liu, Z. (2016). *Exploring cross-linguistic influence: Perception and production of L1 , L2 and L3 bilabial stops by Mandarin Chinese speakers*. Universitat Autònoma de Barcelona.

Mennen, I. (2015). Beyond Segments: Towards a L2 Intonation Learning Theory. 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). Springer. https://doi.org/10.1007/978-3-662-45168-7_9

Prieto, P., & Roseano, P. (2018). Prosody: Stress, Rhythm, and Intonation. In K. L. Geeslin (Ed.), *The Cambridge Handbook of Spanish Linguistics* (pp. 211–236). Cambridge University Press. https://doi.org/10.1017/9781316779194.011

Racine, I., & Detey, S. (2019). Production of French close rounded vowels by Spanish learners. In M. Gibson & J. Gil (Eds.), *Romance Phonetics and Phonology* (pp. 381–394). Oxford University Press. https://doi.org/10.1093/oso/9780198739401.003.0019

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

Shao, J. (2014). *Xiandai Hanyu Yiwenju Yanjiu [Interrogative Sentences in Modern Chinese]*. The Commercial Press.

Sóskuthy, M. (2021). Evaluating generalised additive mixed modelling strategies for dynamic speech analysis. *Journal of Phonetics*, *84*. https://doi.org/10.1016/j.wocn.2020.101017

Turner, J. (2023). Phonetic Development of an L2 Vowel System and Tandem Drift in the L1: A Residence Abroad and L1 Re-Immersion Study. *Language and Speech*, *66*(3), 756–785. https://doi.org/10.1177/00238309221133100

van Rij, J., Wieling, M., Baayen, R. H., & Rijn, H. van. (2022). *itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs*.

Vanrell Bosch, M. del M., Feldhausen, I., & Astruc, L. (2018). The Discourse Completion Task in Romance prosody research: Status quo and outlook. In I. Feldhausen, J. Fliessbach, & M. del M. Vanrell Bosch (Eds.), *Methods in prosody: A Romance language perspective* (pp. 191–227). Language Science Press.

Wieling, M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics*, *70*, 86–116. https://doi.org/10.1016/j.wocn.2018.03.002

Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R* (2nd ed.). Chapman and Hall/CRC.

Xi, X., & Li, P. (2022). Spanish stops and their allophones produced by proficient Mandarin learners of Spanish. In V. H. Do, C. M. Luong, S. Nakamura, H. H. Nguyen, & T. M. H. Nguyen (Eds.), *25th Conference of the Oriental COCOSDA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSDA)* (pp. 152–157). IEEE. https://doi.org/10.1109/O-COCOSDA202257103.2022.9997932

Xi, X., & Li, P. (2023). Effects of stress and prominence on Spanish stops and lenition in L2 speech of proficient Mandarin learners of Spanish. In R. Skarnitzl & J. Volín (Eds.), *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS2023)* (pp. 2740–2744). Guarant International.

Xi, X., Li, P., Baills, F., & Prieto, P. (2020). Hand gestures facilitate the acquisition of novel phonemic contrasts when they appropriately mimic target phonetic features. *Journal of Speech, Language, and Hearing Research*, *63*(11), 3571–3585. https://doi.org/10.1044/2020_JSLHR-20-00084

Xi, X., Li, P., & Prieto, P. (2023). Reducing acoustic overlap of L2 English vowels through gestures encoding lip aperture. In A. Kirkova-Naskova & A. Henderson (Eds.), *Proceedings of the 7th International Conference on English Pronunciation: Issues and Practices* (pp. 261–270). Université Grenoble-Alpes. https://doi.org/10.5281/zenodo.8225191

Xi, X., Li, P., & Prieto, P. (2024). Improving Second Language Vowel Production With Hand Gestures Encoding Visible Articulation: Evidence From Picture-Naming and Paragraph-Reading Tasks. *Language Learning*. https://doi.org/10.1111/lang.12647

Xi, X., Zhou, S., & Li, P. (2024). Duration as a prosodic marker of contextual factors in Mandarin positive polar questions. In Y. Chen, A. Chen, & A. Arvaniti (Eds.), *Proc. Speech Prosody 2024* (pp. 687–691). https://doi.org/10.21437/SpeechProsody.2024-13

Zerbian, S., Downing, L. J., & Kügler, F. (2009). Introduction: Tone and intonation from a typological perspective. *Lingua*, *119*(6), 817–826. https://doi.org/10.1016/j.lingua.2007.10.024

Zhao, T. (2019). La entonación de las preguntas del español hablado por chinos. *Phonica*, *15*, 119–140. https://doi.org/10.1344/phonica.2019.15.119-140

Zhao, T. (2023). Rasgos melódicos de la entonación del español hablado por chinos [Ph.D. Thesis, Universitat de Barcelona]. http://hdl.handle.net/10803/689035

Zhao, T., & Font Rotchés, D. (2022). Interrogativas y entonación de español por los chinos. In B. Blecua Falgueras, J. Cicres i Bosch, M. Espejel, & M. J. Machuca (Eds.), *Propuestas en fonética experimental: enfoques metodológicos y nuevas tecnologías* (pp. 318–323). Universitat de Girona, Servei de Publicacions.