The Intonational Phonology of Requests in L2 Spanish

Robert Esposito

Kendra V. Dickinson

# Author note

Correspondence concerning this article should be addressed to Robert Esposito, Rutgers University - Department of Spanish and Portuguese, 15 Seminary Place, New Brunswick, NJ 08904, USA . E-mail: rme70@rutgers.edu.

# Abstract

This is where the abstract will go.

*Keywords:* Pitch, Prosody, Intonation, Intonational Phonology, Pragmatics, Linguistics

*Word count:* WORD\_COUNT

# 1. The Intonational Phonology of Requests in L2 Spanish

# 2. Introduction

Despite the fact that all languages modulate F0 to encode pragmatic meaning (Arvaniti & Fletcher, 2020), the L2 classroom typically lacks explicit content on intonation. When encountering the L2 outside of the classroom, the lack of instruction on intonation could leave learners disadvantaged in perception (Casillas et al., 2023) and production (Herrero & Devı́s, 2020). Intonation provides one avenue for modulating politeness (Brown & Levinson, 1987), without which could lead to miscommunication when intonation-politeness mappings do not match between the L1 and L2 (Estebas-Vilaplana, 2014).

Outside of the L2 classroom, intonation in modern linguistics is only recently receiving more attention. Where previous models treated intonation with the goals of speech synthesis (Hart et al., 1990) or pedagogy (Pike, 1945), Ladd (2008) and colleagues spearheaded the field of intonational phonology and developed the Autosegmental Metrical theory of intonational phonology (henceforth AM) from which to study intonation from a phonological perspective. The vast majority of research on intonation investigates monolingual speakers, but there do exist recent works that investigate the acquisition of L2 intonation (Astruc & Mar Vanrell, 2016; Nibert, 2005), as well as attempt to model L2 intonation (Mennen, 2015). Other researchers attempt to find universals among the uses of pitch, such as finding a relationship between higher pitch and greater politeness (Brown & Levinson, 1974; Ohala, 1983).

The current research undertaken here investigates the L2 intonation system by looking at L1 English L2 Spanish intonational strategies to express varying levels of politeness. L2 Spanish utterances will be annotated using ToBI within AM framework, and pitch contours will be compared with expected L1 Spanish nuclear configurations. Furthermore, the mean pitch of the utterance and the height of the final boundary tone will be investigated as a function of politeness to see if there is support for a universal association between high pitch and politeness.

## 2.1 Politeness

Brown & Levinson (1987) provide a landmark investigation in linguistic expressions of politeness. In their conception, politeness strategies are used to mitigate face threatening acts, which can attack an individual’s negative face, one’s basic claim to freedom of action and from imposition, or positive face, the self-image one has of oneself and the desire for it to be appreciated and approved of. Politeness strategies are similarly divided into negative politeness strategies, used to minimize the effects of an inevitable face threatening act, and positive politeness strategies, where an individual attempts to align themselves closer with the interlocutor and demonstrate that they share their desires. For example, negative politeness strategies may include being pessimistic about the interlocutor’s response to a request, using questions, and using more complex syntax to hedge (e.g., in Spanish, the imperfect subjunctive); positive politeness strategies, on the other hand, may include using in-group identity markers, seeking agreement on safe topics, or presupposing common ground.

The decision to opt for a specific politeness strategy depends on a cost-benefit analysis of three binary factors: power, social distance, and level of imposition of request. These three independent variables, Brown & Levinson (1987) claim, are assigned different values cross-culturally, which account for the rich diversity in cultures globally.

Power can come from both material and metaphysical sources. Naturally, certain conceptions of power can have more sway in some contexts than others. Monetary power in a bargaining situation would hold much more sway than in a situation where a student is asking for help on an essay from a writing center.

Social distance is characterized by how socially close two individuals are. For example, the social distance between two siblings are typically much lower than the social distance between two students who don’t know each other. Of note, a student and a professor who have a close relationship may have a very low social distance, but an imbalance of power. Significantly, social distance here does not refer directly to the distance assigned to them based on their station, but to their personal relationship.

Finally, the level of imposition of request transparently refers to how imposing the request is. For example, asking someone to borrow a pencil is a low imposition in nearly all situations; however, asking anyone at all to borrow their expensive laptop proves to be a much higher imposition.

In the current study, the three identified factors will be closely controlled for in the context of the United States university environment, appealing to the target population of college students. The specific situations will be defined below.

## 2.2 Autosegmental Metrical (AM) Theory

Within the AM framework (Ladd, 2008), the continuous phonetic cue F0 is mapped categorically to ordered strings of Low (L) and High (H) tones that attach to stressed syllables, called pitch accents, and edges of phrases, called boundary tones. Significantly, the phonological representations are not meant to include all elements of pitch modulation, just as all phonetic elements are not recorded for segments. That is, phonetically distinct contours may be realization of the very same underlying phonological element. Thus, the goal of AM is not to record every phonetic detail of F0, but to understand how F0 maps to discrete, categorical tones that express meaning (Ladd, 2022).

The abstract, phonological representation of intonation with the AM framework is carried out through the Tones and Breaks Indices (ToBI) tool, originally developed for General American English (Silverman et al., 1992). This tool has also been developed for Spanish, called Sp\_ToBI (Beckman et al., 2002). Of note is that ToBI requires that each variety of a language have its own ToBI defined, based on its individual intonational inventory. For example, various varieties of English (Grice et al., 2020; Mayo et al., 1997; Silverman et al., 1992) and Spanish (Prieto & Roseano, 2010) have since been documented using ToBI.

[Figure 1](#fig-tobi-example) gives an example of an utterance labeled using ToBI. ToBI associates a sound file with its spectrogram and pitch track, as well as several tiers of intonation. The typical structure is a transcription tier, typically divided into syllables and represented with either the language’s orthography or IPA, a tone tier, where pitch accents and boundary tones are aligned with their associated syllables, and optionally a break index tier, representing the perceived strength of prosodic boundaries.

|  |
| --- |
| Figure 1: Spectrogram and pitch track (blue) of the utterance “Ana lleva el abrigo” (“Ana is wearing a coat”) labeled using ToBI. |

On the tone tier, pitch accents are represented with **\*** (e.g., H\* represents a High tone associated with a stressed syllable), while **%** is associated with boundary tones (e.g. H% is a High boundary tone at the end of a phrase). A **-** is used to represent a phrase boundary that is not utterance final (e.g., H- is a High phrase boundary). As mentioned above, tones can be combined to represent rises and falls, and when used as pitch accents, the **\*** indicates the tone that is most associated with the stressed syllable. For example, H\*L is a fall where the stressed syllable is affiliated with the H tone; HL\* is a fall where the stressed syllable is affiliated with the L tone. Crucially, each ToBI system will define how an underlying representation can be realized, such that “H\*L” may be realized as the phonetically differently in two different languages or two different varieties of the same language.

It is commonly found that there is a distinction between “prenuclear” and “nuclear” pitch accents, where the nuclear pitch accent is the rightmost pitch accent, and all those that come before it are prenuclear. Typically, studies documenting intonation in the AM framework find that there are restrictions as to which tones can occur in prenuclear vs nuclear position, but overall there is overlap. Documentary work involves determining which pitch accents can occur in which positions for which utterance types ([Figure 2](#fig-pitch-accent-example)).

|  |
| --- |
| Figure 2: An example of the pitch accent L+H\* in Castilian Spanish with its phonetic realization (stressed syllable in grey) and distribution of use (Aguilar et al., 2009). |

Despite the previous claim that intonation is mapped to categorical tones, there is evidence that this is not entirely true. Ladd & Morton (1997), for example, suggests that there is a distinction to be made between production and perception: whereas production is, in fact, categorically determined, perception is gradient. Orrico & D’imperio (2020) supports the view that perception of intonation is gradient by investigating epistemic bias (i.e., expecting a positive or negative response) in the perception of yes-no questions in Solerno Italian. In their experiment, L1 Solerno Italian speakers listened to yes-no questions. The questions’ nuclear pitch accent peak and boundary tone peak were synthetically manipulated in various height steps. Participants were asked to choose a point on a scale from 0 “the speaker expects NO” to 100 “the speaker expects YES”. They found that participants with higher empathy *and* experience with other Italian varieties had a more granular perception of the steps, whereas other participants had a more binary perception.

Although there is clearly evidence for a gradient approach to perception of intonation, there is less research that investigates production. This study will help contribute to this body of literature by investigating the height of boundary tones in the production of L1 English L2 Spanish learners to indicate politeness. Under the frequency code (Ohala, 1983), which will be explained below,it’s predicted that higher levels of politeness would be associated with a higher boundary tone pitch. That is, in this study, where power, social distance, and level of imposition of the request are set to [+], we would expect the highest boundary tones, and for the boundary tone to become gradiently lower as the independent variables are step-wise set to [-].

## 2.3 L2 Intonation

AM has been primarily developed to investigate L1 intonation, but there have been recent attempts to expand intonation investigation into the L2 domain. For example, Mennen (2015) has proposed the theoretical model L2 Intonation Learning Theory (LILt). Four dimensions of intonation are recognized under this model with the goal of characterizing the similarities and differences between the intonation inventories between two languages:

1. The inventory and distribution of categorical phonological elements (“systemic” dimension)
2. The phonetic implementation of these categorical elements (“realizational” dimension)
3. The functionality of the categorical elements of tunes (“semantic” dimension)
4. The frequency of use of the categorical elements (“frequency” dimension)

The LILt model, as demonstrated by the specified dimensions, are primarily concerned with L1-L2 transfer effects. That is, it is predicted that a language learner acquiring an L2 will experience positive transfer effects when the intonation systems align along one or more dimensions, whereas they will experience negative transfer when they do not align. For example, English and Spanish differ in how they express polite responses to wh-questions. Where English uses a broad pitch range in its rises, represented with by the nuclear configuration L+¡H\* (where **¡** represents an extra high tone), Spanish uses a narrow pitch range, represented by L+H\* (Estebas-Vilaplana, 2014).

The differences here have a significant impact on the acquisition of Spanish by L1 English speakers. English speakers have available to them L+H\* (Dilley & Heffner, 2013), the expected contour in Spanish (“systemtic” dimension). It is realized phonetically similarly, although the English L+H\* typically has a broader pitch range (for English: Beckman & Elam, 1997; for Spanish: Prieto & Roseano, 2010) (“realizational” dimension”). The “frequency” dimension is defined at a more general level by Mennen (2015) and would require more space than attributed here to investigate the frequency differences of tone primitives in English and Spanish. However, crucially, there is a mismatch in the *semantic* dimension. Where English uses L+¡H\* as an appropriate, polite response to a wh-question, Spanish does not pattern in this way; instead, it is mapped to an “over-excited” response, which could give rise to misinterpretation when learners interact with monolingual L1 Spanish speakers.

The intonation of Spanish in L1 English L2 Spanish learners will be investigated under the basic tenets of AM, using ToBI to document their underlying contours, as well as work to test the LILt model’s predictions on L1-L2 intonation transfer. Broadly, it is expected that where General American English and Spanish match along the aforementioned dimensions, L1 English L2 Spanish learners will experience positive transfer; where there are mismatches along any of the dimensions, they will experience negative transfer. That is, we expect to see typical L1 English contours in their L2 Spanish.

## 2.4 Pitch and Politeness

Although AM takes for granted the arbitrary nature of sound to meaning correspondences that dominate human language, some researchers propose that some aspects of pitch-meaning correspondences are non-arbitrary. For example, Ohala (1983), in his “frequency code”, the use of pitch in human language evolved from a cross-species use of F0, where higher pitch is “nonthreatening, submissive, subordinate, in need of the receiver’s cooperation and good will”, whereas lower pitch is “aggressive, assertive…” (p. 8). He claims that these correspondences are cross-linguistic, citing the universality of F0’s use to distinguish sentence modality. That is, cross-linguistically, low pitch is associated with statements (“assertiveness”), whereas high pitch is associated with questions (“in need of the receiver’s cooperation and good will”). Under this belief, using higher pitch is a cross-linguistic method of expressing politeness.

This belief is supported by Brown & Levinson (1987). They give Tzeltal as an example, where politeness is marked by high pitch or falsetto. They claim elsewhere that there is a universal association between high pitch and tentativeness (Brown & Levinson, 1974).

Given the cited research, it would be expected that some of the earliest politeness strategies adopted by L2 learners would involve pitch, as opposed to (morpho-)syntax or semantics, as they already have the correspondence between higher pitch and politeness available to them in their L1. This is in opposition to more syntactically complex expressions of politeness strategies, such as using the imperfect subjunctive in Spanish.

Importantly, as demonstrated by Estebas-Vilaplana (2014), a broader pitch range in a given *pitch accent* is not always positively associated with politeness. As such, this, along with the frequency code (Ohala, 1983), gives motivation to examine pitch more broadly along two dimensions: the overall mean pitch of an utterance and the height of the final boundary tone of the utterance.

# 3. The Present Study

Few studies have worked to document the intonation of L2 Spanish (Henriksen et al., 2010; Nibert, 2005), and even fewer still have investigated the intersection of intonation and politeness Astruc & Mar Vanrell (2016). The present study aims to document the most common nuclear pitch accents and boundary tones for requests produced by L1 English L2 Spanish speakers, which will then be examined through the lens of the LILt model (Mennen, 2015). Furthermore, this study aims to put the frequency code to the test (Ohala, 1983) by examining pitch height (at the utterance-level and at the final boundary tone) as a function of politeness.

# 4. Previous Research

Spanish intonation is rarely explicitly taught in the L2 classroom. However, this seems a disservice to L2 learners, as evidenced by the possible miscommunications evinced by Estebas-Vilaplana (2014) (see above for analysis). The possible miscommunications due to intonation were further explored by Herrero & Devı́s (2020). In their study, L1 Mandarin Chinese L2 Spanish speakers produced polite requests in Spanish that were then rated by L1 Spanish speakers on a 6-point Likert scale, where 0 was most impolite and 5 was most polite. It was found that even when other pragmatic strategies were used to mitigate impoliteness, such as lexical features (e.g., *gracias* “thank you”), the speakers were still perceived overall as impolite. Herrero & Devı́s (2020) claim that the perceived impoliteness was due to the use of zig zag patterns and falling contours, which are described as impolite melodic patterns in Spanish (Devı́s, 2011). The present study hopes to continue pushing for the inclusion of intonation in the L2 classroom by demonstrating the miscommunications that may occur if L2 learners do not acquire intonation.

The only study that specifically investigates L1 English L2 Spanish politeness intonation production is Astruc & Mar Vanrell (2016). In their study, they created two Spanish-language corpora: one comprised of L1 Mexican Spanish speakers and one comprised of L1 UK English L2 Mexican Spanish speakers. The L2 corpus was comprised of 28 utterances, 14 invitations and 14 requests, produced by 14 A2 learners. These were produced as a homework assignment for their class, where the assignment was to “leave a telephone message for a friend telling him or her what they did last night, suggesting one thing that they could do together at the weekend, and requesting that this friend returned their call” (p. 7). In this situation, power, social distance, and level of imposition of request were all set to [-], as they are speaking to a friend ([-power, -social distance]) and imposing a small imposition ([-imposition]).

In their analysis of L2 corpus’s intonation, Astruc & Mar Vanrell (2016) found that requests were most associated with the low rise, high rise, simple fall, and fall rise, the last of which was exclusively associated with requests. They found influence from English in the choice of intonation patterns, such as the overuse of falling intonation patterns (corresponding to the “frequency” dimension described by Mennen (2015)). Most importantly, they found that the overall distribution and frequency of each tonal accent differed from the L1 corpus: high rises were much more frequent in the L2 corpus, as well as the number of falling patterns being much more diverse. They also produced utterances with a broader pitch range, which would possibly be interpreted as over-excited by L1 Spanish speakers (Estebas-Vilaplana, 2014).

Although we do not expect to see the same intonation contours as Astruc & Mar Vanrell (2016) due to dialect differences (they studied UK English speakers, the present study investigates Americans), we do expect to see L1-L2 transfer, as they did, as well as a broader pitch excursion as compared to monolingual L1 Spanish speakers.

## 4.1 Research Questions

Our first research question is exploratory in nature, whereas our third research question is experimental.

1. What pitch accents do L1 General American English L2 Spanish learners use for requests in Spanish?

Based on Astruc & Mar Vanrell (2016), we expect there to be L1-L2 transfer of pitch contours, such that L1 English pitch contours will be used in the L2 Spanish.

1. Do pitch accents map categorically to the independent variables power, social distance, and level of imposition of request (Brown & Levinson, 1987)?

Due to insufficient previous data in this domain, we must rely on L1 research to inform our hypothesis. Astruc & Mar Vanrell (2016) reported that in their L1 corpus, only the level of imposition of the request presented a categorical usage of pitch accent, and so we expect here similar results: power and social distance will not categorically map to pitch accents, but level of imposition of request will.

1. Does higher pitch correlate with higher politeness (Ohala, 1983)?

This question will be examined quantitatively in two ways. First, the overall mean pitch of the utterance will be examined, and then the height of the final boundary tone. These two measures have been chosen due to their salience in the speech stream.

# 5. Methods

The experimental stimuli were designed as a discourse completion task (DCT) that controlled for power, social distance, and level of imposition (Brown & Levinson, 1987), all of which had two steps for a total of 8 conditions [Table 1](#tbl-conditions). Each condition had 3 items. Experimental stimuli were written by the first author and reviewed by a native speaker of Chilean Spanish who has experience teaching Spanish to the target population.

Table 1: Conditions for experimental stimuli.

| Condition | Power | Distance | Level of Imposition |
| --- | --- | --- | --- |
| 1 | - | - | - |
| 2 | - | - | + |
| 3 | + | + | - |
| 4 | + | + | + |
| 5 | + | - | - |
| 6 | + | - | + |
| 7 | - | + | - |
| 8 | - | + | + |

# 6. Scraps

Due to the limited data available to them, Astruc & Mar Vanrell (2016) called for studies investigating a range of proficiency levels. This study hopes to fill that gap by documenting higher proficiency L1 English L2 Spanish learners. Notably, this study will investigate American English college-aged students, as opposed to the British English speakers, so the contours found in their study from the L2 speakers are not directly comparable to the ones found here.

This study will investigate polite requests in L1 English L2 Spanish learners through a discourse completion task (DCT).

Fundamental frequency (F0), and its phonological correlate pitch, is used cross-linguistically for a variety of functions. While some languages modulate F0 at both lexical and post-lexical levels, some languages, such as English and Spanish, primarily modulate F0 at the post-lexical level. At this level, we refer to the modulation as “intonation”, in which case it encodes pragmatic meaning and phrase boundaries.

Given these proposals, it would be expected that prosodic features may be some of the first politeness strategies positively transferred from the L1 to the L2, due to the crosslinguistic similarities and ready availability. A variety of politeness strategies require more complex morpho-syntactic transformations (cite brown here), but something as simple as using higher pitch to express increasing politeness is something available to even the most neophyte language learner. This study aims to investigate the intonational resources that L2 Spanish speakers have available to them.

All languages make use of F0, and its phonological correlate pitch, at the post-lexical (phrasal) level to encode phrasal boundaries and pragmatic meaning. We refer to this usage as “intonation” (Arvaniti & Fletcher, 2020). For example, intonation in some languages can be used to differentiate sentence modality or expressing doubt (Ladd, 2008); another use, under investigation here, is to express politeness (Brown & Levinson, 1987).

# References

Aguilar, L., De-la-Mota, C., & Prieto, P. (2009). *Sp\_ToBI training materials*. <http://prosodia.upf.edu/sp_tobi/>. <http://prosodia.upf.edu/sp_tobi/>

Arvaniti, A., & Fletcher, J. (2020). *The autosegmental-metrical theory of intonational phonology*.

Astruc, L., & Mar Vanrell, M. del. (2016). Intonational phonology and politeness in L1 and L2 spanish. *Probus*, *28*(1), 91–118.

Beckman, M. E., Dı́az-Campos, M., McGory, J. T., & Morgan, T. A. (2002). *Intonation across spanish, in the tones and break indices framework*.

Beckman, M. E., & Elam, G. A. (1997). Guidelines for ToBI labelling, version 3. *Ohio State University*.

Brown, P., & Levinson, S. (1974). A theory of prosodic semantics. *Unpublished MS. Language Behavior Research Lab, Univ. Of California, Berkley*.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge university press.

Casillas, J. V., Garrido-Pozú, J. J., Parrish, K., Arroyo, L. F., Rodrı́guez, N., Esposito, R., Chang, I., Gómez, K., Constantin-Dureci, G., Shao, J., et al. (2023). Using intonation to disambiguate meaning: The role of empathy and proficiency in L2 perceptual development. *Applied Psycholinguistics*, *44*(5), 913–940.

Devı́s, E. (2011). La entonación de (des) cortesı́a en el español coloquial. *Phonica, 2011, Num. 7, p. 36-79*.

Dilley, L. C., & Heffner, C. C. (2013). The role of F0 alignment in distinguishing intonation categories: Evidence from american english. *Journal of Speech Sciences*, *3*(1), 3–67.

Estebas-Vilaplana, E. (2014). The evaluation of intonation. *Evaluation in Context. Philadelphia: John Benjamins*, 179–194.

Grice, M., German, J. S., & Warren, P. (2020). Intonation systems across varieties of english. *The Oxford Handbook of Language Prosody*.

Hart, J., Collier, R., & Cohen, A. (1990). A perceptual study of intonation. *(No Title)*.

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

Herrero, C., & Devı́s, E. (2020). Unintentional impolite intonation in L2 spanish requests produced by chinese workers living in madrid. *Proceedings of the 10th International Conference on Speech Prosody*, 2020–2173.

Ladd, D. R. (2008). *Intonational phonology*. Cambridge University Press.

Ladd, D. R. (2022). *The trouble with ToBI*.

Ladd, D. R., & Morton, R. (1997). The perception of intonational emphasis: Continuous or categorical? *Journal of Phonetics*, *25*(3), 313–342.

Mayo, C., Aylett, M., & Ladd, D. R. (1997). Prosodic transcription of glasgow english: An evaluation study of GlaToBI. In *Intonation: Theory, models and applications*.

Mennen, I. (2015). Beyond segments: Towards a L2 intonation learning theory. In *Prosody and language in contact: L2 acquisition, attrition and languages in multilingual situations* (pp. 171–188). Springer.

Nibert, H. J. (2005). The acquisition of the phrase accent by intermediate and advanced adult learners of spanish as a second language. *Selected Proceedings of the 6th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*, 108–122.

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

Orrico, R., & D’imperio, M. (2020). Individual empathy levels affect gradual intonation-meaning mapping: The case of biased questions in salerno italian. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *11*(1), 1–39.

Pike, K. L. (1945). *The intonation of american english.*

Prieto, P., & Roseano, P. (2010). *Transcription of intonation of the spanish language*. Lincom Europa Munich.

Silverman, K. E., Beckman, M. E., Pitrelli, J. F., Ostendorf, M., Wightman, C. W., Price, P., Pierrehumbert, J. B., & Hirschberg, J. (1992). ToBI: A standard for labeling english prosody. *ICSLP*, *2*, 867–870.