# Are you asking? An ERP Study on the Perception of Early Intonational Cues in L1 and L2 Spanish

Robert Esposito

Department of Spanish and Portuguese, Rutgers University

# Author Note

Thank you to Nuria Sagarra, who introduced me to neurolinguistics and the importance of on-line research methods.

Correspondence concerning this article should be addressed to Robert Esposito, Department of Spanish and Portuguese, Rutgers University, 15 Seminary Place, New Brunswick, NJ 8904, USA, Email: rme70@rutgers.edu

# Abstract

Accurately parsing intonation is crucial for identifying sentence modality (i.e., declarative versus interrogative utterances) in Spanish. For example, *Mariano habla del tiempo* may variably be a statement (‘Mariano talks about the weather’) or a question (‘Does Mariano talk about the weather?’) depending on the intonation. Traditional analysis of Spanish intonation points to solely the final pitch movement as determining sentence modality ([Tomás Navarro, 1944](#ref-tomas1944manual)), but a recent gating experiment has demonstrated that L1 Spanish speakers can identify sentence modality by the first word ([Face, 2007](#ref-face2007role)). The Predictive Coding model ([Van Petten & Luka, 2012](#ref-van2012prediction)) theorizes that humans integrate linguistic information incrementally to make predictions about future events. As such, it follows that Spanish speakers would use early intonational cues to predict sentence modality, instead of waiting for the final pitch movement. This study uses electroencephalogram (EEG) to investigate the sensitivity to early intonational information for identifying sentence modality in L1 and L2 Spanish individuals of varying proficiency levels. Participants completed a two-alternative forced choice task in which they identified if an utterance was a question or not. In the target items, participants were exposed to utterances in which the first pitch peak was manipulated to be higher (for base-declarative utterances) or lower (for base-interrogative utterances). The findings of the present study will inform the Predictive Coding model by further elucidating what linguistic content individuals integrate as sentence processing unfolds.

*Keywords*: ERP, EEG, prosody, intonation, suprasegmental, interrogatives, perception

# Are you asking? An ERP Study on the Perception of Early Intonational Cues in L1 and L2 Spanish

As early as infancy, humans are sensitive to fundamental frequency (F0), the acoustic correlate of intonation. In perception, newborns are sensitive to the F0 signal in singing and continuous speech ([Sambeth et al., 2008](#ref-sambeth2008sleeping)); in production, newborn cry melodies are shaped by the intonation of their native language ([Prochnow et al., 2019](#ref-prochnow2019does)). As early as 5 months old, infants acquiring European Portuguese can distinguish between statements and yes-no questions, which are different only in intonation ([Frota et al., 2014](#ref-frota2014infants)). Both native speakers (L1 speakers) and second language learners (L2 speakers) must accurately produce or perceive intonation to prevent communicative breakdowns. For example, distinguishing a sarcastic exclamation and a genuinely complimentary one in English (e.g., “How smart she is!”) requires correct parsing of intonation, as the utterance is identical at both the lexical (word) and syntactic (word-order) levels ([Zhou et al., 2019](#ref-zhou2019detecting)).

Even more quotidian, Spanish declaratives and yes-no questions (known as absolute interrogatives) are similarly distinguished only at the level of intonation: *Mariano habla del tiempo* “Mariano talks about the weather” versus *¿Mariano habla del tiempo?* “Does Mariano talk about the weather?” Traditionally, only the final F0 movement of the utterance is analyzed as providing the disambiguating “key” to determine if such an utterance is a declarative or interrogative ([Tomás Navarro, 1944](#ref-tomas1944manual)). However, Face ([2007](#ref-face2007role)) has determined in an off-line gating experiment that L1 Spanish speakers can distinguish between these two sentence types as early as the first word, which encompasses the first pitch movement. While the final pitch movement strongly influenced participants’ judgement, it appears that L1 Spanish speakers can use early intonational cues to determine sentence modality (i.e., whether an utterance is a declarative or interrogative), even before the final movement is reached.

Although early intonation cues are available to L1 and L2 Spanish speakers to predict sentence modality, it is unclear if they are processed to make predictions about upcoming linguistic content. Theoretical models of brain function suggest that the brain makes continuous predictions about future events based on Bayesian inferences ([Bar, 2007](#ref-bar2007proactive); [Lee & Mumford, 2003](#ref-lee2003hierarchical)), giving rise to the Bayesian Brain theory ([Knill & Pouget, 2004](#ref-knill2004bayesian)). One instantiation of the Bayesian Brain theory is the Predictive Coding model ([Friston, 2005](#ref-friston2005theory)), which has been proposed to underlie language comprehension ([Hagoort & Indefrey, 2014](#ref-hagoort2014neurobiology); [Huettig, 2015](#ref-huettig2015four); [Van Petten & Luka, 2012](#ref-van2012prediction)). Within the Predictive Coding model, cognition works to minimize surprise by making accurate predictions with minimal error to reduce processing time. Under this framework, it would be expected that L1 Spanish and “proficient” (that is, native-like) L2 Spanish speakers would integrate intonational information as sentence processing unfolds to predict sentence modality.

The present study aims to determine if L1 and L2 Spanish speakers are sensitive to the early intonational cues identified by Face ([2007](#ref-face2007role)) during on-line sentence processing using electroencephalogram (EEG) while participants complete a two-alternative forced choice task (FCT). Specifically, participants will be tasked with identifying if an aurally presented utterance is a question. Importantly, the first pitch movement will be synthetically modified to conflict with the base utterance’s sentence modality (e.g., a declarative base sentence with a question-like first pitch movement). If participants are sensitive to and integrate the first pitch movement as processing unfolds to predict sentence modality, it is expected that the subsequent pitch movements would elicit the P3a and P3b components, as it provides task-relevant conflicting information about the sentence’s modality.

# Background

This study investigates how intonation is integrated into sentence processing, guided by the Predictive Coding model. The following section outlines the theoretical foundation of the Predictive Coding model, emphasizing its application to language. Intonation’s role in language is then discussed, highlighting the Autosegmental Metrical framework. The processing of intonation in native speakers and second language learners is then reviewed.

## The Predictive Coding Model

With the advent of neuroimagining methods such as the electroencephalogram (EEG), neuroscientists have actively sought to understand mechanisms underlying brain functions. One theory is that the brain proactively makes predictions about the future ([Bar, 2007](#ref-bar2007proactive)), and researchers have developed the concept of the Bayesian Brain ([Knill & Pouget, 2004](#ref-knill2004bayesian)). According to the Bayesian Brain, human brains make inferences about the causes of sensations based on a generative model of the world ([Friston, 2012](#ref-friston2012history)). A specific instantiation of this is the Predictive Coding (PC) model, which was originally developed for understanding visual processing ([Rao & Ballard, 1999](#ref-rao1999predictive)). Under this model, cognition works to minimize surprise by making accurate predictions, which reduces processing time for events.

Although PC was originally used to model visual processing, it has since been extended to account for language processing as well ([Van Petten & Luka, 2012](#ref-van2012prediction)). Through methods that allow for observation of on-line sentence processing, such as EEG and eye tracking, researchers have determined that many aspects of the speech stream are integrated incrementally and immediately, as opposed to all at once at the end. This can be seen by the famous N400 effect. In the landmark study be Kutas and Hillyard ([1980](#ref-kutas1980reading)), participants silently read sentences while being recorded with EEG. Participants were exposed to two conditions: semantically correct sentences (e.g., “It was his first day at work”) and sentences where the last word was semantically incongruent with the context (e.g., “He spread the warm bread with *socks*”). The semantically incongruent condition elicited an N400 effect, a negative-going wave between 250 ms and 550 ms and peaking at 400 ms after stimulus onset. The N400 effect is now typically thought to reflect the brain’s attempt to integrate incoming information with prior expectations, reflecting greater processing effort for unexpected linguistic content ([Kutas & Federmeier, 2011](#ref-kutas2011thirty)).

It is still not entirely clear what aspects of intonation are integrated as sentence processing unfolds, but on-line research studies have demonstrated that some aspects of intonation are integrated to make predictions. For example, Esteve-Gibert et al. ([2020](#ref-esteve2020empathy)) investigated the integration of intonational information in a visual world paradigm with L1 French speakers. Participants listened to an interaction between two speakers playing a card game, where Speaker A had to guess the image on Speaker B’s card. The target items were homophones, such as *cane* “female duck/walking stick”. If Speaker A guessed correctly, Speaker B responded, “I have a female duck, indeed, the animal”; if Speaker A guessed incorrectly, Speaker B responded, “I have a walking stick, instead, for walking”. Crucially, Speaker B’s card could be predicted before the homophone was disambiguated in the second part of the sentence by an intonational cue, named the “contradiction contour”, in the first part of the sentence. Participants with high empathy, measured with the empathy quotient ([Baron-Cohen & Wheelwright, 2004](#ref-baron2004empathy)), recruited this intonational information to predict the correct semantic meaning of the homophone. The results indicate that, at least some individuals, integrate intonational cues as sentence processing unfolds, in line with PC.

The current study aims to understand what other aspects of intonation are integrated to make predictions during speech perception. Where Esteve-Gibert et al. ([2020](#ref-esteve2020empathy)) focused on how intonational cues like the “contradiction contour” may guide semantic disambiguation in sentence processing, the present research investigates how intonational cues are integrated to distinguish a Spanish declarative and an interrogative (i.e., sentence modality). Traditionally, only the final F0 movement of a Spanish utterance is thought to disambiguate a declarative and an interrogative, but evidence from Face ([2007](#ref-face2007role)) and predictions from PC make it seem more likely that other intonational cues earlier in the sentence are integrated to predict sentence modality. What follows is a review of intonation and the major framework through which intonation is studied, the Autosegmental Metrical theory of intonation.

## Intonation and the Autosegmental Metrical Framework

Intonation has often been relegated to a secondary role in a language’s grammar. This is particularly surprising, as intonation is one of the first linguistic features perceived by humans cross-linguistically, beginning as early as the 20th week of gestation when the auditory system becomes operational ([Gervain, 2018](#ref-gervain2018gateway)). Intonational phonologists have pushed for recognizing intonation as equally central to a language as its syntax or semantics, in part due to its universal status: all languages make use of intonation to encode pragmatic meaning. Broadly speaking, intonation can encode speech act marking, information status, belief status, and politeness, among other functions ([Arvaniti & Fletcher, 2020](#ref-arvaniti2020autosegmental); [Best, 2019](#ref-best2019diversity); [Prieto, 2015](#ref-prieto2015intonational)).

More concretely, and most relevant to the present study, intonation can encode sentence modality (i.e., whether an utterance is a declarative or interrogative). In Spanish, intonation has such a function: compare, for example, *Ya comió* “He already ate” and *¿Ya comió?* “Did he eat already?”. These two utterances differ primarily in the modulation of fundamental freqeuncy (F0), which is determined by how fast the vocal folds vibrate. The continuous F0 signal is perceived in human language as pitch (e.g., a “high” pitch or a “low” pitch), and pitch is then organized into patterns, giving rise to intonation. Traditionally, the final F0 movement in Spanish determines sentence modality: a falling F0 marks a declarative, whereas a rising F0 marks an interrogative ([Tomás Navarro, 1944](#ref-tomas1944manual)).

Currently, the Autosegmental Metrical (AM) theory of intonation is the leading framework for studying intonation ([Ladd, 2008](#ref-ladd2008intonational); [Pierrehumbert, 1980](#ref-pierrehumbert1980submitted)). Under this model, the continuous F0 contour maps to phonologically discrete, sequential tone targets. Tone targets can be H(igh), L(ow), or a combination of the two to represent rises (LH) or falls (HL), which are organized as either pitch accents, anchored to the stressed syllable, or edge tones anchored to prosodic boundaries. Edge tones are furthered divided into phrase accents and boundary tones, anchored to non-terminal prosodic boundaries and the Intonation Phrase boundary, respectively. Pitch accents are indicated with \* (e.g., H\*); non-final edge tones with - (e.g., H-); and boundary tones with % (e.g., H%). These representations and phonetic realizations are part of a many-to-many relationship, such that a single representation, such as H+L\*, may have various phonetic realizations both within and across languages. This parallels segmental phonology, where a single segment, such as the Spanish /b/, may have various realizations depending on numerous contextual factors, such as a voiced bilabial plosive, bilabial approximant, or voiced labio-dental fricative.

Furthermore, the final pitch accent (deemed the nuclear pitch accent) combined with the boundary tone receive a special status as the nuclear configuration, which is typically mapped to sentence type ([Prieto & Roseano, 2010](#ref-prieto2010transcription)). However, Pierrehumbert ([1980](#ref-pierrehumbert1980submitted)) proposed that prenuclear pitch accents (i.e., pitch accents that occur to the left of the nuclear pitch accent) also contribute information to the sentence. The current study investigates the integration of prenuclear pitch accents during on-line sentence processing and prediction of sentence modality, which will address an important gap in our understanding of on-line sentence processing of early intonational information, further contextualized below, by investigating the temporal dynamics of intonation processing.

## L1 Intonation Processing

Although some areas of language have been well-investigated using neurocognitive methodologies, such as morphosyntax, there remains a large gap concerning intonation. However, what studies have been completed have repeatedly demonstrated the importance of intonation, supporting claims by intonational phonologists that intonation is a critical component of a language’s grammar. For example, Li ([2018](#ref-li2018temporal)) investigated the hierarchical processing of intonation using EEG in Mandarin Chinese by comparing responses to F0 changes produced by intonation or lexical tone in two forced choice tasks. In the first “intonation task”, participants identified whether the sentence was a statement or question; in the second “tone task”, participants identified the tone (tone 2 or tone 4) of the final word. Li ([2018](#ref-li2018temporal)) found that participants processed intonation before lexical tone in both tasks: between 150-250 ms in the first task after the target word, and beginning at 100 ms in the second. These results seem to indicate the priority that intonation has in sentence processing.

Beyond simply the temporal hierarchy of intonation processing, researchers have investigated the effects of inappropriate or missing pitch accents, relevant to the present study. Such anomalies have been found to elicit a range of ERP components, including the N400, P3a, and P3b ([Ito & Garnsey, 2004](#ref-ito2004brain); [Johnson et al., 2003](#ref-johnson2003erp); [Liu et al., 2016](#ref-liu2016online); [Magne et al., 2005](#ref-magne2005line)). Due to a lack of research on Spanish using EEG methodologies, the present study must be informed by ERP components elicited by anomalous pitch accents in other languages. The ERP components focused on in this study are the P3a and P3b subcomponents, which comprise part of the P300 family of components, also known as the Late Positive Component ([Polich, 2007](#ref-polich2007updating)). These components are a positive-going peak, typically found within the time window 250-500 ms. The P3a is associated with novel and surprising events, while the P3b is associated with surprising events that are task-relevant ([Donchin & Coles, 1988](#ref-donchin1988p300)).

These specific ERP components have been chosen due to previous research on French, a closely related language to Spanish. Magne et al. ([2005](#ref-magne2005line)) investigated how anomalous pitch accents affect on-line sentence processing in L1 French. Participants listened to question-answer dialogues where the lexical (word) and syntactic (word-order) content of the answers remained identical, but the pitch accents varied. For example, in response to a question like, “Did he give his fiancée a ring or a bracelet?”, the corresponding answer, “He gave a ring to his fiancée” should have a pitch accent on *ring*, not *fiancée*. Participants judged whether the responses were coherent with the preceding question. Magne et al. ([2005](#ref-magne2005line)) found that incongruous pitch accents in sentence-medial positions (e.g., *ring*) elicited the P3a and P3b components, while those in sentence-final position (e.g., on *fiancée*) elicited an N400 effect. Since the present study focuses on sentence-medial pitch accents, the P3a and P3b components are those under investigation.

While these findings provide valuable insight into how native speakers process intonation, much less is known about how L2 learners process similar cues. For example, are L2 Spanish speakers sensitive to secondary cues indicating sentence modality? The following section addresses the theoretical framework and literature that investigates L2 intonation processing.

## L2 Intonation Processing

The insights gained from L1 intonation processing studies provide a foundation for understanding how intonation is processed in L2 contexts. While frameworks like AM theory have been extensively applied to model L1 intonation, there is less work done on L2 intonation. This is particularly important, as we know that even highly proficient L2 learners may rely on different cognitive strategies when processing their second language ([Caffarra et al., 2017](#ref-caffarra2017end)). Understanding how L2 learners process intonation in real time is crucial for developing and improving models of L2 acquisition and for designing effective language teaching methods.

Although other areas of language processing have been addressed in research, such as L2 morphosyntactic processing, few researchers have attempted to address how intonation is processed in the L2. Mennen ([2015](#ref-mennen2015beyond)) proposed models such as the L2 Intonation Learning Theory (LILt), which primarily highlights transfer effects between the L1 and L2 intonational inventories in four dimensions: the inventory and distribution of categorical phonological elements (systemic), the phonetic implementation of these categorical elements (realizational), the functionality of the categorical elements (semantic), and the frequency of use of the categorical elements (frequency). However, focusing on solely transfer effects ignores the individual differences between learners, such as proficiency, working memory, and language experience, as well as the general cognitive burden of processing an L2.

Some researchers have attempted to address these individual differences specifically in the domain of intonation. For example, Casillas et al. ([2023](#ref-casillas2023using)) presented a conceptual replication of Brandl et al. ([2020](#ref-brandl2020development)) with L1 English L2 Spanish participants. Participants completed the LexTALE-ESP ([Izura et al., 2014](#ref-izura2014lextale)), a measure of Spanish proficiency, and the empathy quotient ([Baron-Cohen & Wheelwright, 2004](#ref-baron2004empathy)), a measure of empathy. They then completed an off-line two-alternative forced choice task, where they were exposed to aural utterances and had to determine if they were a question. Casillas et al. ([2023](#ref-casillas2023using)) found an interaction between proficiency and empathy, such that lower proficiency individuals with higher empathy more accurately identified sentence modality than their lower empathy counterparts. Furthermore, as would be expected, accuracy increased with high proficiency. Such a finding is not accounted for by the LILt model ([Mennen, 2015](#ref-mennen2015beyond)).

While foundational frameworks like AM theory and the LILt model have provided valuable insights into the processing of intonation in L1 and L2 contexts, they fall short of addressing individual differences that influence L2 intonation development, acquisition, and processing. Recent research highlights the importance of factors such as proficiency and empathy, which shape a learner’s ability to interpret intonational cues in their L2. These findings indicate that a more comprehensive model is necessary for addressing L2 intonation, as transfer effects are not sufficient. The current research seeks to address this gap.

## Background Summary

In summary, the present study hopes to contribute to our understanding of the processing of intonation in L1 Spanish speakers, as well as the development, acquisition, and processing of intonation in L2 Spanish speakers. Guided by the Predictive Coding model, it would be expected that speakers recruit early intonational information to make predictions about sentence modality, as opposed to waiting for the final cue at the end of the utterance. Furthermore, the current models for L1 and L2 intonation do not account for cognitive factors that influence processing and development. The following section reviews the specifics of the present study and theoretical motivations based on Face ([2007](#ref-face2007role)).

# The Present Study

Traditional analysis of Spanish intonation posits that the final boundary tone determines if a sentence is a declarative (L%) or an absolute interrogative (H%) ([Tomás Navarro, 1944](#ref-tomas1944manual)). However, Face ([2007](#ref-face2007role)) has found that there exist other intonational cues earlier in the speech stream that indicate sentence modality. In a behavioral gating experiment, Face ([2007](#ref-face2007role)) determined that L1 Spanish speakers can accurately determine sentence modality as early as the first pitch accent (i.e., the first full word) of an utterance due to the higher pitch peak in interrogatives. Although the final boundary tone was found to be the most informative cue for sentence modality, almost deterministic in nature, it appears that native speakers can make use of the first pitch peak as a secondary cue to determine sentence modality.

With respect to Spanish learners, Bedialauneta Txurruka ([2023](#ref-bedialauneta2023perception)) investigated secondary cues to determine sentence modality in a two-alternative force task, where participants had to identify if an auditory stimulus was a question or not. Although she did not report proficiency for L2 participants, all L2 participants had over 5 years of Spanish learning experience. She found that participants relied on universal tendencies to interrogation, primarily using the final boundary tone to identify sentence modality, whereas native speakers variably integrated other cues to determine sentence modality.

Although intonational information in the speech stream before the boundary tone are evidently not deterministic for deciding sentence modality ([Bedialauneta Txurruka, 2023](#ref-bedialauneta2023perception); [Face, 2007](#ref-face2007role)), native Spanish speakers and Spanish learners may still be sensitive to them during on-line sentence processing. The present study investigates if Spanish natives and learners are sensitive to early intonational cues during on-line sentence processing while determining if a sentence is a declarative or absolute interrogative. This study builds on the previous work by reporting proficiency information for L2 participants.

The three research questions for this study are experimental in nature. The first research question address L1 Spanish speakers, whereas the second and third research questions address L2 Spanish speakers.

1. Are L1 Spanish speakers sensitive to intonational cues before the utterance-final boundary tone to predict a given utterance’s sentence modality?

Contrary to traditional analysis of Spanish intonation ([Tomás Navarro, 1944](#ref-tomas1944manual)), L1 Spanish speakers seem to be able to use intonational information before the final boundary tone to determine a sentence’s modality ([Bedialauneta Txurruka, 2023](#ref-bedialauneta2023perception); [Face, 2007](#ref-face2007role)). However, it is unclear if native speakers use these cues as sentence processing unfolds to make predictions. Given Predictive Coding theory, which predicts that humans constantly update predictions based on incoming information, it would be expected that L1 speakers process the early intonational information to predict sentence modality. This hypothesis would be supported quantitatively if incongruous prenuclear pitch accents elicit a P3a or P3b in the forced choice task presented to them.

1. Are L2 Spanish speakers sensitive to intonational cues before the utterance-final boundary tone to predict a given utterance’s sentence modality?

L2 speakers have a greater cognitive load when processing their L2. Due to the difficulty that Spanish learners have when identifying sentence modality ([Brandl et al., 2020](#ref-brandl2020development); [Casillas et al., 2023](#ref-casillas2023using)), I predict that learners will not universally show sensitivity to early intonational information to predict sentence modality, and no P3a or P3b will be elicited by incongruous prenuclear pitch accents.

1. Is sensitivity to early intonational cues modulated by proficiency?

It has been seen that L2 proficiency modulates behavioral accuracy when identifying sentence type ([Brandl et al., 2020](#ref-brandl2020development); [Casillas et al., 2023](#ref-casillas2023using)). This study will provide cognitive evidence for the role of proficiency on the sensitivity to secondary intonation features during on-line sentence processing by Spanish learners. It is predicted that high proficiency Spanish learners will be able to integrate more intonational information in the speech stream during on-line sentence processing, such that they will pay more attention to the first pitch accent to determine if an utterance is a broad focus declarative or absolute interrogative. Consequently, the mismatch between the incongruous prenuclear pitch accents will elicit P3a and P3b responses, similar to L1 speakers.

# Methodology

The present study investigated L1 and L2 Spanish speakers. Participants completed a forced choice task (FCT) while monitored with EEG. L2 participants also completed two proficiency assessments. The following sections detail participant inclusion criteria, the tasks completed, and the creation of stimuli.

## Participants

40 participants (20 L1 Spanish and 20 L2 Spanish), with an age range of 18 to 28 were recruited to partake in the study. All participants included in this study were right-handed. No participants reported a history of learning, neurophysiological, or hearing disabilities.

L1 Spanish speakers were defined as those who began to acquire Spanish at birth and were raised in a Spanish speaking country. Participants did not spend more than two months in a non-Spanish speaking country. Participants who reported proficiency in another language at any level above A2 in the CEFR framework (i.e., low proficiency) were excluded.

For the L2 Spanish speakers, late bilingualism, in this study, was operationalized as individuals who acquired Spanish after the age of 13. Participants were recruited from college Spanish classes. Participants who spoke a language other than English in their household, or who self-reported above A1 proficiency (functionally monolingual) in another language other than English and Spanish, were excluded. Furthermore, participants who reported having spent more than one month in a Spanish-speaking country were excluded. Participant proficiency ranged from intermediate to high (i.e., native-like), as we were interested in how the acquisition of secondary intonational cues developed in a range of proficiencies.

## Tasks

All data were collected individually in one session (approximately 1.5 hours). L2 speakers first completed two Spanish proficiency assessments before the experiment (15 min), and then completed the two-alternative forced choice task while monitored using EEG (1 hour). L1 speakers only completed the two-alternative forced choice task while monitored using EEG.

### Proficiency Assessment

L2 participants first completed two Spanish proficiency assessments. The first test was an abbreviated version of the *Diploma de Español como Lengua Extranjera* (DELE), which has been used in previous studies (e.g., [Sagarra & Herschensohn, 2010](#ref-sagarra2010role)). The test consists of 56 multiple choice questions that assess grammar and vocabulary knowledge. Correct answers receive one point, and incorrect answers receive zero. L2 participants then completed the Lexical Test for Advanced Learners of Spanish (LexTALE-ESP, henceforth LexTALE) ([Izura et al., 2014](#ref-izura2014lextale)). The LexTALE is a lexical decision experiment that provides a standardized assessment of vocabulary size in Spanish. Participants are exposed to a series of real and pseudowords on the computer screen, and they must decide if the word is real or fake. LexTALE scores can range from -20 to 60. Monolingual Spanish speakers generally score above 50, whereas individuals with little or no knowledge tend to score negative. Adult learners with low to intermediate proficiency range from 0 to 25, and advanced learners generally score above 25. Only L2 participants who scored at an intermediate proficiency range or higher were included in this study.

### Two-Alternative Forced Choice Task

For the two-alternative forced choice task, participants identified if an aurally presented utterance was a question or not using the keyboard. The left and right shift keys were used and the answer assigned to each key (either “yes, this is a question” or “no, this is not a question”) was pseudo-randomized for each participant. Participants completed one of two versions of the task, assigned pseudo-randomly, which included 20 items from each condition. The items were pseudo-randomized for each participant. Participants did not encounter any lexically/syntactically identical items, such that if a participant was exposed to “Ana lleva el abrigo” in Condition 1, they were not exposed to the same utterance in any other condition. Participants were exposed to 40 experimental and 40 control items, for a total of 80 items.

The stimuli were recorded by a native speaker of Castilian Spanish. The speaker produced syntactically and lexically identical utterances, once produced as a broad focus declarative ([Figure 1](#fig-original-d)) and again as an information-seeking absolute interrogative ([Figure 2](#fig-original-q)). The utterances were then manipulated using the stylize pitch feature in Praat ([Boersma & Weenink, 2024](#ref-boersma2024praat)). For the manipulated stimuli, the first pitch accent was manipulated to raise or lower the peak for the declarative ([Figure 3](#fig-manipulated-d)) or interrogative ([Figure 4](#fig-manipulated-q)), respectively. The modified pitch peak heights were decided based on the corresponding utterance’s pitch peak height.

Figure 1

Original spectrogram with pitch contour for the broad focus declarative, Ana lleva el abrigo. ‘Ana is wearing the coat.’

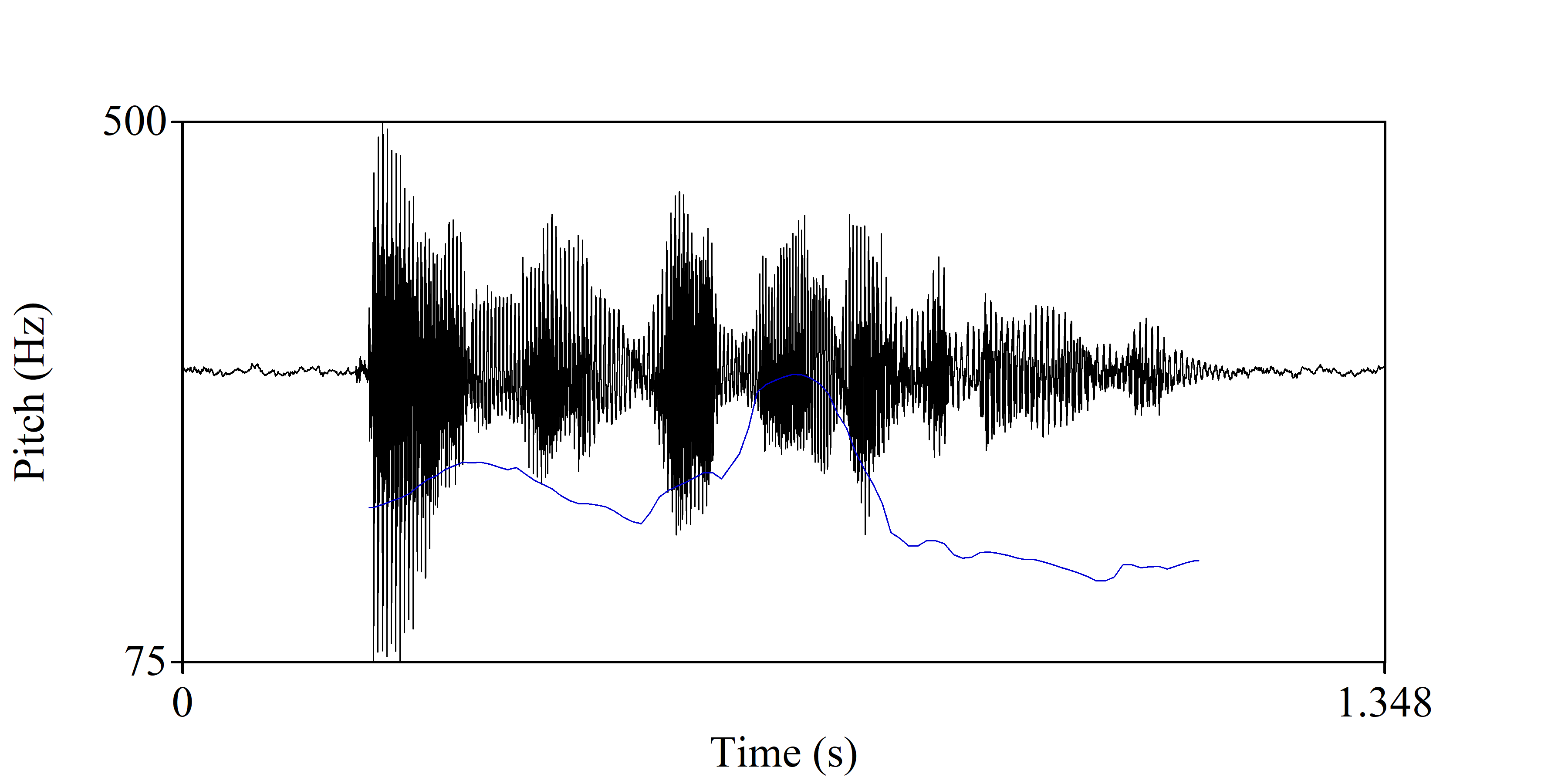


Figure 2

Original spectrogram with pitch contour for the information-seeking absolute interrogative, ¿Ana lleva el abrigo? ‘Is Ana wearing the coat?’

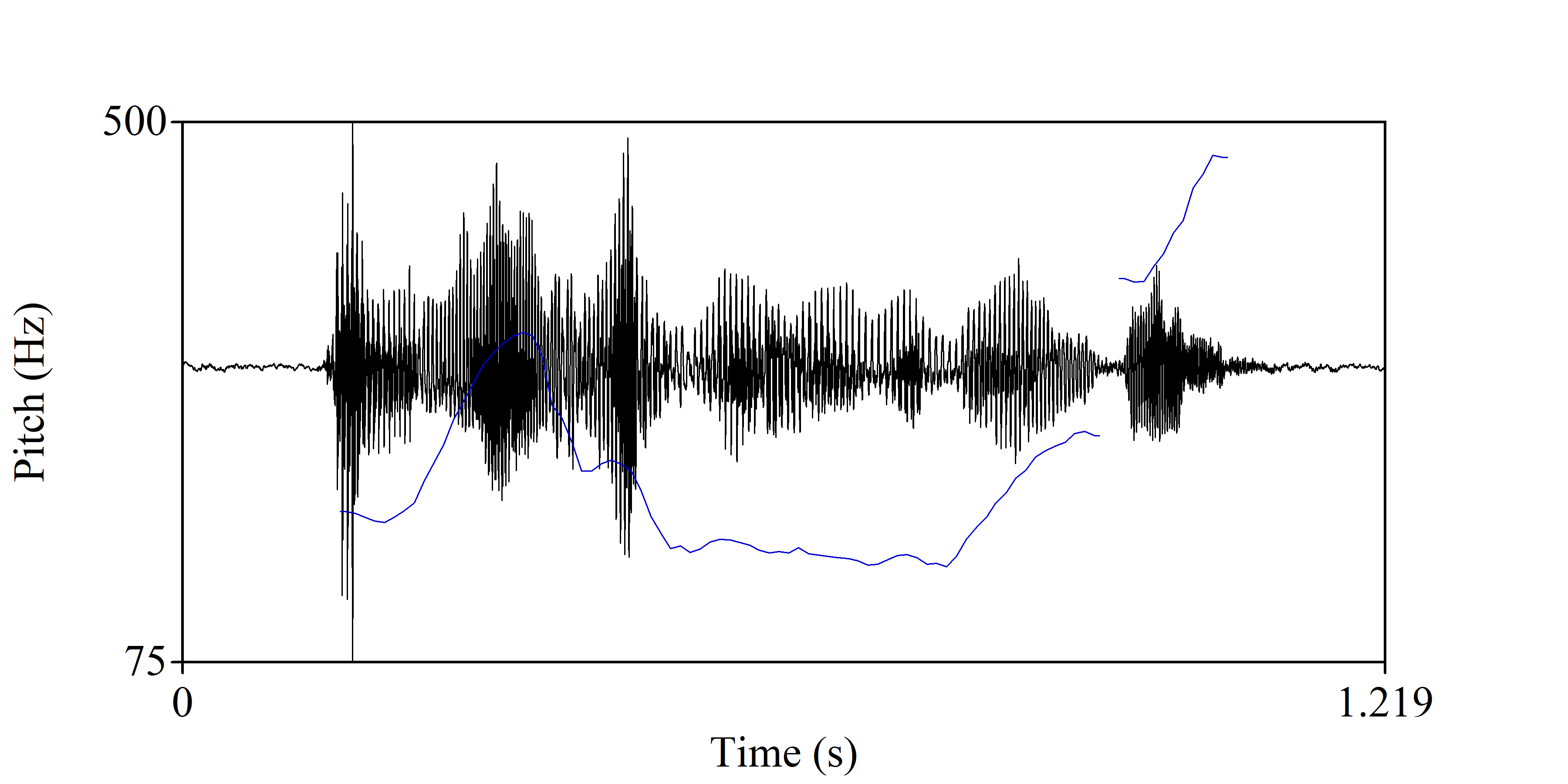


Figure 3

Manipulated pitch contour for the utterance Ana lleva el abrigo. ‘Ana is wearing the coat.’

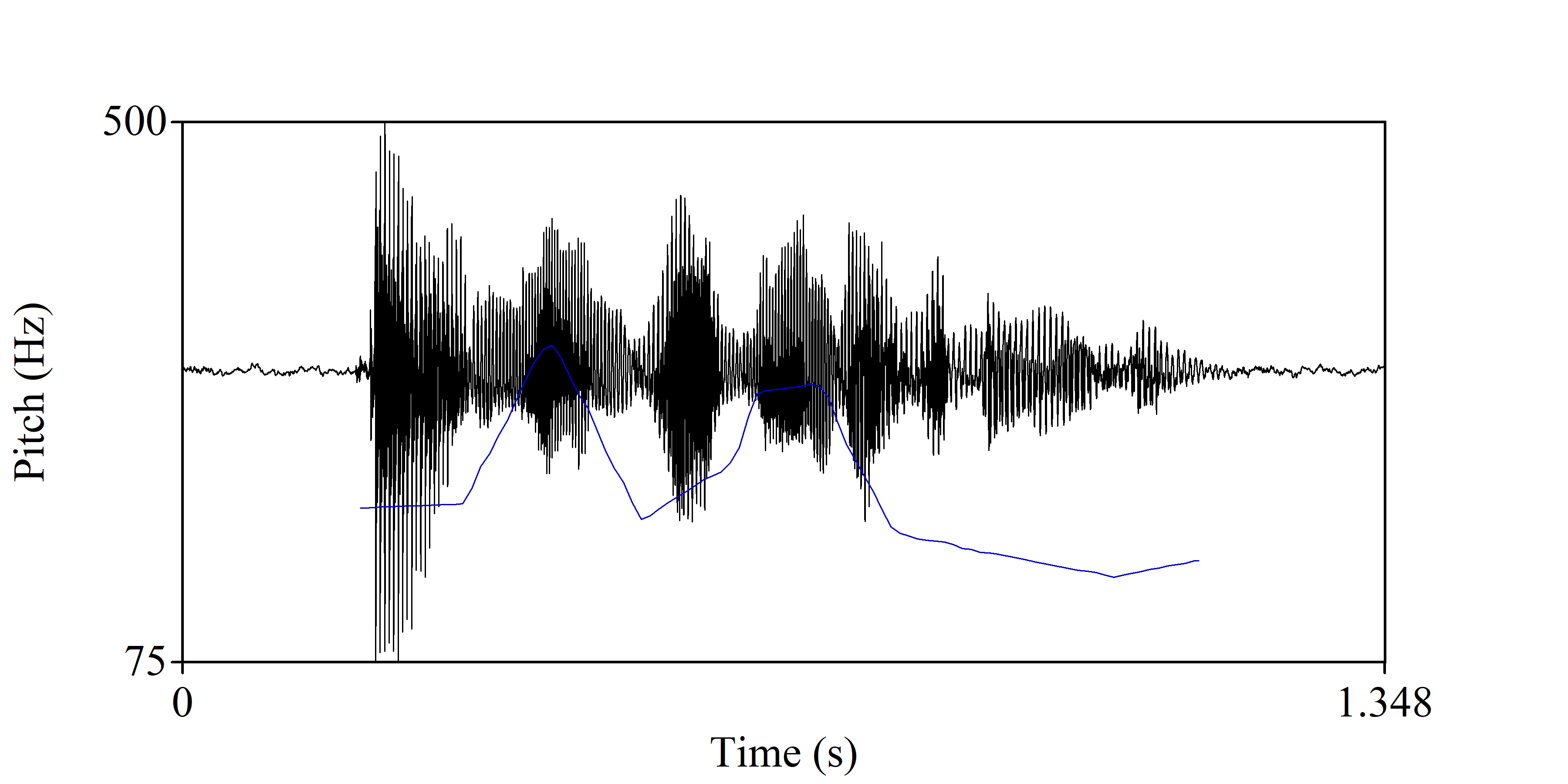
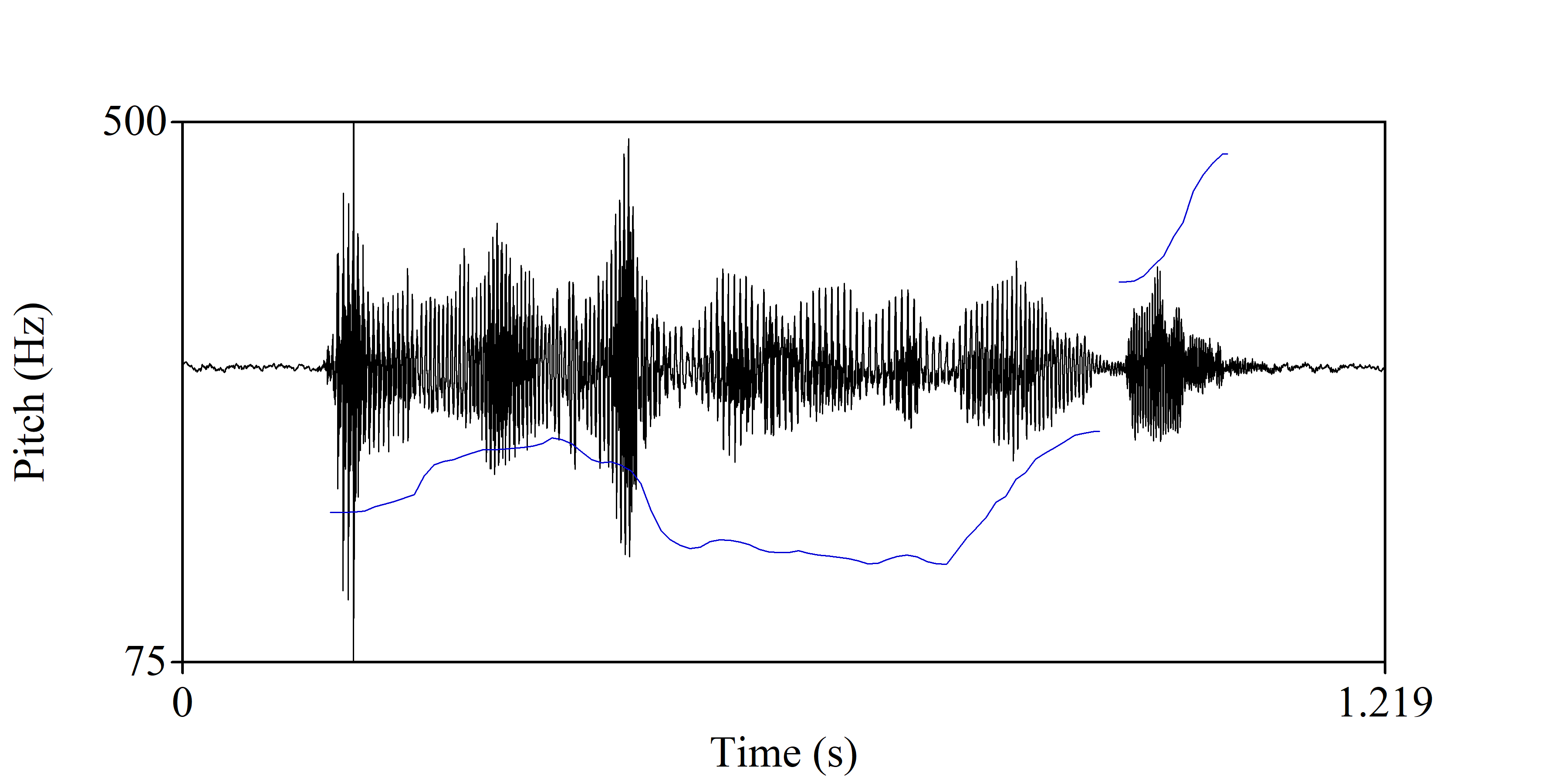


Figure 4

Manipulated pitch contour for the utterance ¿Ana lleva el abrigo? ‘Is Ana wearing the coat?’



In total, there were four conditions in a 2X2 design: (1) the original declarative, (2) the base declarative utterance with the first pitch accent peak raised, (3) the original interrogative, and (4) the interrogative utterance with the first pitch accent peak lowered. All conditions can be seen in [Table 1](#tbl-conditions). Each condition had 80 items, for a total of 320 items.

Table 1

Conditions for the forced-choice task.

| Condition | Modality | First Pitch Accent Peak |
| --- | --- | --- |
| 1 | declarative | original |
| 2 | declarative | lowered |
| 3 | interrogative | original |
| 4 | interrogative | heightened |

## Procedure

Subjects were seated in a comfortable chair located in a sound and electrically attenuated room. They were instructed to keep their eyes fixated on a fixation cross on the monitor and to avoid movement as much as possible, including eye movements, blinks, and jaw clenching, as well as to remain as relaxed as possible. Acoustic stimuli were presented via *computer information* through *headphone information*.

Participants were familiarized with the task through the same 32 training items. They were instructed to decide if an utterance was a question or not by using the keyboard (left and right shift keys) as accurately and as quickly as possible once the audio stimulus ended. They were signaled to respond once the audio ended with a question mark “?” that appeared on the screen. They did not receive feedback for accuracy during the training or during the experimental items. After the training, participants were randomly assigned Version 1 or Version 2 of the task.

During the training and experimental items, participants were instructed to fixate on a fixation cross on the computer screen. The fixation cross remained on the screen during the entirety of the audio stimulus, and as soon as the audio stimulus finished, a question mark “?” appeared on the screen, signalling for them to decide if the utterance was a question or not. The interval between stimuli varied between 2.5 and 5.0 s. Participants were given a 3-5 minute break after the 32 training items, and another break after 40 items, or when requested.

# References

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

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, *11*(7), 280–289.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *34*, 163–175.

Bedialauneta Txurruka, I. (2023). Perception of spanish declarative questions and statements by L2 spanish speakers. In R. Skarnitzl & J. Volín (Eds.), *Proceedings of the 20th international congress of phonetic sciences* (pp. 2591–2596). Guarant International.

Best, C. T. (2019). The diversity of tone languages and the roles of pitch variation in non-tone languages: Considerations for tone perception research. *Frontiers in Psychology*, *10*, 364.

Boersma, P., & Weenink, D. (2024). *Praat: Doing phonetics by computer [computer program]*. <http://www.praat.org/>

Brandl, A., González, C., & Bustin, A. (2020). The development of intonation in L2 spanish. *Hispanic Linguistics: Current Issues and New Directions*, *26*, 11.

Caffarra, S., Barber, H., Molinaro, N., & Carreiras, M. (2017). When the end matters: Influence of gender cues during agreement computation in bilinguals. *Language, Cognition and Neuroscience*, *32*(9), 1069–1085.

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.

Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*(3), 357–374.

Esteve-Gibert, N., Schafer, A. J., Hemforth, B., Portes, C., Pozniak, C., & D’Imperio, M. (2020). Empathy influences how listeners interpret intonation and meaning when words are ambiguous. *Memory & Cognition*, *48*, 566–580.

Face, T. L. (2007). The role of intonational cues in the perception of declaratives and absolute interrogatives in castilian spanish. *Journal of Experimental Phonetics*, *16*, 185–225.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 815–836.

Friston, K. (2012). The history of the future of the bayesian brain. *NeuroImage*, *62*(2), 1230–1233.

Frota, S., Butler, J., & Vigário, M. (2014). Infants’ perception of intonation: Is it a statement or a question? *Infancy*, *19*(2), 194–213.

Gervain, J. (2018). Gateway to language: The perception of prosody at birth. *Boundaries Crossed, at the Interfaces of Morphosyntax, Phonology, Pragmatics and Semantics*, 373–384.

Hagoort, P., & Indefrey, P. (2014). The neurobiology of language beyond single words. *Annual Review of Neuroscience*, *37*(1), 347–362.

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, *1626*, 118–135.

Ito, K., & Garnsey, S. M. (2004). Brain responses to focus-related prosodic mismatch in japanese. *Speech Prosody 2004, International Conference*.

Izura, C., Cuetos, F., & Brysbaert, M. (2014). Lextale-esp: A test to rapidly and efficiently assess the spanish vocabulary size. *Psicológica*, *35*(1), 49–66.

Johnson, S. M., Breen, M., Clifton, C., & Morris, J. (2003). ERP investigation of prosodic and semantic focus. *Poster Presented at Cognitive Neuroscience, New York City*.

Knill, D. C., & Pouget, A. (2004). The bayesian brain: The role of uncertainty in neural coding and computation. *TRENDS in Neurosciences*, *27*(12), 712–719.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*(1), 621–647.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205.

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

Lee, T. S., & Mumford, D. (2003). Hierarchical bayesian inference in the visual cortex. *JOSA a*, *20*(7), 1434–1448.

Li, J. (2018). *Temporal hierarchy of intonation and tone processing in mandarin chinese*.

Liu, M., Chen, Y., & Schiller, N. O. (2016). Online processing of tone and intonation in mandarin: Evidence from ERPs. *Neuropsychologia*, *91*, 307–317.

Magne, C., Astésano, C., Lacheret-Dujour, A., Morel, M., Alter, K., & Besson, M. (2005). On-line processing of “pop-out” words in spoken french dialogues. *Journal of Cognitive Neuroscience*, *17*(5), 740–756.

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.

Pierrehumbert, J. B. (1980). *Submitted to the department of linguistics and philosophy on 9 september 1980 in partial fulfill ment of the requirements for the degree of doctor of philosophy* [PhD thesis]. MASSACHUSETTS INSTITUTE (OF TECHNOLOGY.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148.

Prieto, P. (2015). Intonational meaning. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(4), 371–381.

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

Prochnow, A., Erlandsson, S., Hesse, V., & Wermke, K. (2019). Does a ‘musical’mother tongue influence cry melodies? A comparative study of swedish and german newborns. *Musicae Scientiae*, *23*(2), 143–156.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87.

Sagarra, N., & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 spanish. *Lingua*, *120*(8), 2022–2039.

Sambeth, A., Ruohio, K., Alku, P., Fellman, V., & Huotilainen, M. (2008). Sleeping newborns extract prosody from continuous speech. *Clinical Neurophysiology*, *119*(2), 332–341.

Tomás Navarro, T. (1944). Manual de entonación española. *Colección Málaga*.

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, *83*(2), 176–190.

Zhou, W., Nolan, F., Li, A., Gong, J., Xue, X., & Ji, X. (2019). Detecting sarcastic and complimentary exclamations in english with intonation patterns and acoustic features: A case study. *2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, 1–6.