Using intonation to disambiguate meaning: The role of empathy and proficiency in L2 perceptual development

Joseph V. Casillas1, Nicole Rodríguez2, Juan José Garrido Pozú1, Kyle Parrish1, Laura Fernández Arroyo1, Robert Esposito3, Isabelle Chang1, Kimberly Gómez1, Gabriela Constantin-Dureci1, Jiawei Shao1, Iván Andreu1, & Katherine Taveras1

1 Rutgers University

2 Poznan

3 Duo-lingo

Author note

We express our gratitude to all the haters.

The authors made the following contributions. Joseph V. Casillas: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing; Nicole Rodríguez: Writing - Review & Editing; Juan José Garrido Pozú: Writing - Review & Editing; Kyle Parrish: Writing - Review & Editing; Laura Fernández Arroyo: Writing - Review & Editing; Robert Esposito: Writing - Review & Editing; Isabelle Chang: Writing - Review & Editing; Kimberly Gómez: Writing - Review & Editing; Gabriela Constantin-Dureci: Writing - Review & Editing; Jiawei Shao: Writing - Review & Editing; Iván Andreu: Writing - Review & Editing; Katherine Taveras: Writing - Review & Editing.

Correspondence concerning this article should be addressed to Joseph V. Casillas, Rutgers University - Department of Spanish and Portuguese, 15 Seminary place, New Brunswick, NJ 08904, USA . E-mail: [joseph.casillas@rutgers.edu](mailto:joseph.casillas@rutgers.edu)

Abstract

The present study investigates the interplay between proficiency and individual pragmatic skills in the process of learning a new language. Notably, we focus on the role of empathy in the development of second language (L2) prosody by analyzing the perception and processing of intonation in questions and statements in L2 Spanish. It is common for L2 learners to struggle with L2 intonation, often resulting in comprehension and communication difficulties (Trofimovich & Baker, 2006). Previous research attests that learners gradually acquire target-language prosody as they gain proficiency in the language. Concretely, the perception and processing of L2 intonation has been shown to improve in conjunction with proficiency conditional on intonation type (Brandl, González, & Bustin, 2020), with polar (‘yes/no’) interrogatives being more difficult to process and acquire when compared with simple statements. The construct empathy has been shown to influence native language processing in how listeners interpret intonation and meaning when words are ambiguous (Esteve-Gibert et al., 2020). Importantly, higher empathy individuals, in comparison with lower empathy individuals, appear to be more sensitive to intonation cues in the process of forming sound-meaning associations. We extend this research to L2 acquisition in order to determine if individual differences in pragmatic skills affect the development of intonation in L2 processing and sentence comprehension. A total of N adult L2 Spanish learners (L1 English) from the Northeastern United States completed a two-alternative forced choice (2AFC) task in which they listened to four utterance types and categorized them as either questions or statements. The stimuli were randomly drawn tokens of declarative (broad, narrow focus) and interrogative (polar, wh-) sentences, spoken by native speakers of eight distinct varieties of Spanish (Andalusian, Argentine, Castilian, Chilean, Cuban, Mexican, Peruvian, Puerto Rican). The stimuli were presented aurally to the participants and subsequently identified as questions or statements using a keyboard. Additionally, participants completed the LexTALE vocabulary task in Spanish (Izura, Cuetos, & Brysbaert, 2014), which served as a proxy for L2 proficiency, as well as the Empathy Quotient questionnaire in English (Baron-Cohen & Wheelwright, 2004), which provided an individual assessment of the construct empathy. We used Bayesian multilevel regression and Drift Diffusion models to analyze the 2AFC data as a function of proficiency and empathy scores. Proficiency and empathy were used as continuous predictors to assess their relationship with pragmatic skill. The results replicated findings from Brandl et al. (2020) showing that learner response accuracy improved as a function of proficiency for all utterance types. Importantly, higher empathy scores were positively correlated with higher accuracy in identifying polar interrogatives. As is the case with L1 research, the present project underscores the importance of considering individual pragmatic differences when examining intonational meaning processing and sentence comprehension in an L2. More notably, the results also motivate the inclusion of measures of pragmatic skill, such as empathy, as predictors for L2 acquisition outcomes. Furthermore, these findings highlight an area in which models of L2 development can improve in order to better account for individual differences in L2 learning.

*Keywords:* Second language acquisition, Intonation, Empathy, Prosody

*Word count:* X

Using intonation to disambiguate meaning: The role of empathy and proficiency in L2 perceptual development

A fundamental difficulty of speech comprehension is that listeners can come to understand different messages when presented with the same linguistic information (Cain, Oakhill, & Lemmon, 2004). This can be especially problematic when one begins the endeavor of learning a new language. In particular, it is common for second language (L2) learners to struggle with intonation—i.e., the melodic contour of an utterance—in the target language (Trofimovich & Baker, 2006). The difficulties associated with intonation can result in comprehension and communication mishaps because the tune is associated not only with linguistic information (e.g., utterance type, syntactic constituency), but also pragmatic information (e.g., polite discourse Astruc, Vanrell, & Prieto, 2016; bias, or presupposition Henriksen, Armstrong, & Garcı́a-Amaya, 2016). The present study investigates how the comprehension of intonation develops in adult L2 learners.

Recent research on monolingual populations suggests that individual differences in pragmatic skills, such as empathy, may play a role in meaning disambiguation (Bishop & Kuo, 2016; Esteve-Gibert, Portes, Schafer, Hemforth, & D’Imperio, 2016; Esteve-Gibert et al., 2020; Orrico & D’Imperio, 2020). Concretely, higher empathy individuals, in comparison with lower empathy individuals, appear to be more sensitive to the intonational cues of speech during the process of forming sound-meaning associations. Furthermore, increasing attention has been given to how individual differences in learner backgrounds play a role in the process of L2 acquisition. The present study contributes to these lines of research by examining how individual differences in pragmatic skills affect the development and processing of intonation during sentence comprehension. Specifically, we investigated the interplay between language proficiency and an individual pragmatic skill (empathy) when learning an L2. We focus on the role of empathy in the development of L2 prosody by analyzing the perception and processing of intonation in questions and statements in L2 Spanish.

## Background and motivation

**L2 acquisition of prosody**. The difficulties associated with learning an additional language in adulthood are numerous. More often than not our focus falls on individual sounds, or segments, though we know that adults who learn an L2 are faced with suprasegmental challenges as well. Concretely, L2 learners often struggle with intonation, i.e., melodic variation at the utterance level. In normal discourse, speakers use intonation to indicate syntactic structure, whether an utterance is a question or a statement, to focus constituents, as well as to convey affective meaning. Notably, the manner in which intonation is mapped to meaning is often language-specific. For these reasons, the development of L2 intonation represents a facet of L2 phonological learning that often results in comprehension and communication difficulties (Trofimovich & Baker, 2006).

As alluded to above, intonation has a semantic function, and through the adequate cognitive decoding of intonation a listener can interpret the function of a given utterance. For example, the intonational contour can indicate whether a speakers’ utterance is declarative or interrogative in nature. Additionally, through prosody a speaker can signal various additional pragmatic functions, such as when they present a polite discourse (Astruc et al., 2016), bias or presupposition (Henriksen et al., 2016). One essential aspect of second language acquisition and language comprehension is that in the presence of the same linguistic elements, listeners can arrive at different interpretations of the message (e.g. Cain et al., 2004). The difficulty in interpreting and decoding the multiple functionalities of intonational contours makes L2 learning a challenge.

Traditionally, intonation is not taught in the L2 classroom. Primary focus is generally on syntax and morphology, with target language phonology receiving much less, if any, attention (REFS). When target language pronunciation is addressed, it often focuses on segmental elements. As a result, intonation is one of the last aspects of L2 phonology that learners acquire (REFS)

Generally, research on L2 intonation has been concerned primarily with production. Learner difficulties are generally ascribed to L1 transfer, and models of L2 phonology focus on the speech segment (i.e., PAM-L2, SLMr, L2LP, though CF trofimovich paper extending SLM) There is a dearth of knowledge with regard to how perception of intonation develops in L2 learning, and even less is known about how individual pragmatic differences account for learner outcomes. The purpose of the present project is to address this gap in the literature by examining the perception and processing of intonation during adult L2 phonological acquisition.

**Acquisition of Spanish prosody**.

Spanish is extensively spoken across the world, with relatively small geolectal differences between its varieties compared to other languages, such that speakers from distinct regions can still generally understand each other. This linguistic continuum allows, however, for variation to occur. Previous research on the acquisition of Spanish prosody attests that learners gradually acquire target-language intonation as they gain proficiency in the language. Research in this area has focused on speech production. For instance, Trimble (2013a) analyzed L2 Spanish learners’ production of intonational patterns for broad focus declaratives and absolute interrogatives after a semester-long study abroad program.

Research on perception of Spanish intonation also supports the notion that mastery is indeed possible for adult learners. Trimble (2013b) examined the perception of intonational cues … Trimble (2013b) found that intonational cues that were absent form participants’ L1 were difficult to perceive. Unsurprisingly, the study suggests the L2 intonation system develops in tandem with proficiency in Spanish, which was positively correlated with time spent studying abroad.

In a similar vein, Brandl et al. (2020) investigated the perceptual development of intonation in questions and statements in L2 Spanish. Specifically, Brandl et al. (2020) examined the effect of L2 proficiency on the perception of statements (broad-focus and narrow-focus) and questions (wh-questions and yes/no questions). Adult English L1-Spanish L2 learners (beginner, intermediate, and advanced) and adult native speakers of Spanish completed an AX discrimination task in which they were presented with two stimuli sentences at a time, one aural and one visual. Participants had to decide whether the sentence presented aurally was the same as the sentence presented visually by pressing a button (‘yes’, ‘no’).

The study found that perception and processing of L2 intonation improved in conjunction with proficiency conditional on intonation type, with polar (‘yes/no’) interrogatives being more difficult to process and acquire when compared with simple statements. Accuracy was high in match conditions, but accuracy in mismatch conditions was below chance for all L2 learners. Regarding response times, intermediate learners were the fastest overall, and native speakers were the slowest. Learners with lower proficiency processed statements faster than questions. Brandl et al. (2020) concluded that L2 intonation perception differs from L1 intonation perception in Spanish, that L2 intonation perception undergoes a gradual development as L2 acquisition progresses, and that statements are more easily processed and acquired by L2 learners of Spanish, compared to yes/no questions.

**Empathy and pragmatic skill**.

The construct empathy has been shown to influence native language processing in how listeners interpret intonation and meaning when words are ambiguous (Esteve-Gibert et al., 2020).

Recent research has also shown that empathy influences native language processing in how listeners interpret intonation and meaning when words are ambiguous (Esteve-Gibert et al., 2020). Specifically, higher empathy individuals, in comparison with lower empathy individuals, appear to be more sensitive to intonation cues in the process of forming sound-meaning associations. In short, individuals with more pragmatic skill (higher empathy) are able to use intonation to resolve temporary lexical ambiguity that can lead to confirmatory vs. contrasting interpretations. This research underscores the importance of considering individual pragmatic differences when examining intonational meaning processing and sentence comprehension. Thus, we extend this research to second language acquisition in order to determine if individual differences in pragmatic skills affect the development of intonation in L2 processing and sentence comprehension.

According to Esteve-Gibert et al. (2020), empathy can be understood as it is related to theory of mind and perspective taking (Baron-Cohen, 2011; Carruthers, 2009; Frith & Frith, 2003). In other words, those with higher empathy are expected to more effectively take the perspective of another, and past research has investigated whether better perspective taking is related to more effective meaning disambiguation. Evidence has been found for the use of intonational cues to disambiguate meaning by higher empathy individuals in monolingual speech (Esteve-Gibert et al., 2020). On the other hand, there is also evidence that higher proficiency Spanish speakers can effectively distinguish questions from statements (Brandl et al., 2020).

Speakers use intonational cues to express feelings, intentions, and pragmatic meaning (see Prieto, 2015), but listeners vary in how they interpret intonational cues to infer meaning (e.g., Bishop & Kuo, 2016, 2016; Portes, Beyssade, Michelas, Marandin, & Champagne-Lavau, 2014). This variability could be related to differences in individual pragmatic skills. Previous studies have shown that individual pragmatic skills correlate with variability in semantic/pragmatic interpretation of ambiguous linguistic items, with more pragmatic individuals preferring pragmatically enriched interpretations and less pragmatic individuals preferring literal/semantic interpretations (e.g., Degen & Tanenhaus, 2016; Nieuwland, Ditman, & Kuperberg, 2010). In addition, more pragmatically skilled individuals and less pragmatically skilled individuals tend to rely on different phonetic cues to parse syntactically ambiguous sentences (Bishop & Kuo, 2016). It is possible that variability in intonation processing is also linked to individual differences in pragmatic skills.

Recent studies have explored the possibility that individual pragmatic skills can modulate the sensitivity of listeners to intonational cues that convey pragmatic meaning. Esteve-Gibert et al. (2020) focused specifically on empathy, which is a pragmatic skill that helps listeners understand other peoples’ feelings, emotions, intentions, and behavior (Baron-Cohen & Wheelwright, 2004). Esteve-Gibert et al. (2020) investigated whether listeners’ empathy modulated variability in intonation processing. Empathy was measured using a self-reported Empathy Quotient (EQ) questionnaire (Baron-Cohen & Wheelwright, 2004). Esteve-Gibert et al. (2020) tested French monolinguals in a visual-world paradigm eye-tracking task that resembled a card guessing game. Participants listened to short dialogues in French between two players (Player A and Player B) playing a card guessing game as they saw a display of four objects on a screen. Participants had to identify the object that depicted the picture of the card of one of the players. Target objects were homophones in French (e.g., cane - “female duck”; canne - “walking stick”). The dialogues consisted of a suggestion by Player A (e.g., “I think you have a stick”) followed by Player B’s response, which could confirm (e.g., “I have a stick, indeed, to walk”) or contrast (e.g., “I have a female duck, instead, the animal”) Player A’s suggestion. The first portion of Player B’s responses varied in intonational contours (confirmation or contrast of beliefs) and had subsequent disambiguating lexical information (indeed/instead + additional information). Results indicated that empathy level affected the processing of an ambiguous word (a homophone) when intonation was the only cue available. Highly empathic individuals varied their looking behavior as a function of intonational cues while less empathic individuals did not. However, the direction of the looking behavior of the highly empathic individuals was not as expected and reverse intonation-meaning mapping was observed. Confirmation intonation did not trigger more looks to the suggested image and contrast intonation did not trigger more looks to objects other than the suggested image.

To explore this unexpected finding further, Esteve-Gibert et al. (2020) tested an additional group of French monolinguals in an offline matching task in which they listened only to the ambiguous portion of Player B’s response (e.g., “I have a stick.”, “I have a female duck.”) and matched the sentences to either a confirmatory or contrastive meaning. Intonation was the only disambiguating cue available. Results showed that the higher the listeners’ empathy, the more accurate they were matching intonation to meaning. Esteve-Gibert et al. (2020) concluded that listeners use empathy as a pragmatic ability to process linguistic information and that empathy influences how listeners use intonation to infer pragmatic meaning, especially when linguistic information is ambiguous or non literal. Pragmatic skills like empathy need to be considered when investigating intonation processing in typical and atypical individuals.

## The present study

We investigate the interplay between proficiency and individual pragmatic skills via empathy in the development of second language (L2) prosody by analyzing the perception of intonation in questions and statements in L2 Spanish.

It is unclear whether those with higher empathy also take advantage of intonational cues to infer another speaker’s communicative intent in a second language when proficiency is also taken into account.

The present study investigates whether L2 learners of Spanish with higher empathy are more effective in distinguishing statements from questions when proficiency is held constant.

the project intends to contribute to cumulative science by conceptually replicating the Brandl et al. (2020) finding, which demonstrates that perception and processing of L2 intonation improves in conjunction with L2 proficiency and as a function of intonation type. Second, this research will extend the current findings on perceptual development to include pragmatic skills, specifically empathy.

This research builds on recent studies looking at the role of individual pragmatic skills in language processing and extends them to the field of second language acquisition. Concretely, we examine the role of empathy in the development of L2 prosody by analyzing the perception of intonation in questions and statements in L2 Spanish. L2 learners gradually acquire Spanish prosody as they gain proficiency in the language. The extant literature shows that ‘yes/no’-type questions are harder to process and acquire, when compared with simple statements (Brandl et al., 2020).

Current study is a conceptual replication of Brandl et al. (2020) - Extend current findings on perceptual development to include pragmatic skills (empathy) - Build on studies that look at the role of individual pragmatic skills in language processing - And if individual differences in pragmatic skills affect the development of intonation in L2 processing and sentence comprehension - Examine the role of empathy in the development of L2 prosody by analyzing the perception of intonation in questions and statements in L2 Spanish

### Research questions.

Research question 1: Is perceptional development in L2 Spanish modulated by proficiency and intonation type (i.e., Brandl et al., 2020)?  
Hypothesis: Accuracy will increase and processing time will decrease as a function of proficiency and intonation type. Yes-no questions will present the most difficulty for L2 learners of Spanish, followed by wh-questions and declarative broad focus and narrow focus statements.

Research question 2: Do pragmatic skills—specifically, empathy—modulate the rate of development in L2 prosody? Hypothesis: Based on the findings of Esteve-Gibert et al. (2020), we posit that prosodic development will occur sooner and at a faster rate in higher empathy individuals. In this operationalization, ‘sooner’ refers to lower proficiency levels in a cross-sectional design, that is, at an earlier developmental stage than lower empathy individuals.

Research question 3: Does speaker variety affect perception accuracy and processing speed? Hypothesis: Based on tentative findings from native speaker pilot data, we hypothesize that overall L2 learners will have most difficulty (lower accuracy, slower RTs) with the Cuban variety.

# Method

## Participants

Participants were adult native speakers of English with varied levels of proficiency in Spanish (determined by the LexTALE task). Participants were be recruited via Prolific and were be compensated at a rate of $9.52 per hour for their time. We estimated the task would take approximately 15 minutes to complete, thus each participant was compensated with $2.70 for completing all three tasks.

Participants were L1 speakers of American English who ranged from functionally monolingual to highly proficient in Spanish (bilinguals). All participants with knowledge of Spanish were adult L2 learners, operationally defined as having begun the endeavor of learning Spanish after the age of 13.

We planed to collect data from 300 individuals (100 monolingual Spanish speakers and 200 L2 learners). Following Brandl et al. (2020), we assumed the effect size for perceptual learning was moderate in terms of the criteria set forth for L2 research by Plonsky and Oswald (2014) (Cohen’s D = 0.600, Pearson’s r = 0.287). Based on this assumption, we estimated that we would need 94 participants to have an 80% chance of capturing the proficiency effect with a type II error rate of 5%. Our hypothesis related to empathy as a possible mediator of perceptual learning is exploratory in nature and thus we did not base our sample size estimate on any parameter estimates related to this effect. That said, we believed the aforementioned exploratory effect was likely to be small, and, considering the resources necessary and available to us, we planned to recruit 100 additional participants.

The pool of online-recruited participants was filtered using criteria set in Prolific.ac to insure participants self-reported as being L1 English speakers born, raised, and currently living in the US with no knowledge of any languages other than English and/or Spanish. They reported no hearing difficulties and were required to use headphones on a desktop computer. Upon beginning the experiment, all participants responded to the following screening questions: 1) What part of the US are you from? 2) At what age did you begin learning Spanish? 3) Are you proficient in any languages other than English/Spanish?. Additionally, participants responded to the prompt “I am most familiar with Spanish from…” and using a pull-down window they selected a variety of Spanish or “I am not familiar with any variety of Spanish”.

We excluded data from any participant that responded that they were not from the US Northeast, that they began learning Spanish before the age of 13, or that they were proficient in a language other than English/Spanish. Participants responding categorically across all trials were also excluded.

## Tasks

The study consists of three tasks, given in the following order: a two-alternative forced choice task (2AFC), followed by the LexTALE, and, finally, the Empathy quotient (EQ).

**2afc**. In the 2AFC task participants will be presented with an audio file containing a statement (declarative: broad focus or narrow focus) or a question (yes-no or wh-). Their task is to determine, as quickly and as accurately as possible, if the utterance they hear is a question or a statement. Specifically, they will respond to the question “Is this a question” using the keyboard. They will type ‘1’ for “yes” (meaning “yes, this is a question”) and ‘0’ for ‘no’ (meaning “no, this is not a question”). The stimuli consist of 64 critical items, 16 of each sentence type. To generate the stimuli, we recorded native Spanish speakers of 8 different varieties (Cuban, Castilian, Andalusian, Puerto Rican, Chilean, Argentinean, Mexican, and Peruvian). The 8 native speakers all produced the same 64 critical items. We segmented all utterances using Praat and normalized them for peak intensity. The 2AFC task consists of 64 trials in which the stimuli presented are randomized across speaker variety. Each variety has the same probability of being selected on a given trial, such that, on average, a given participant will hear each variety approximately 8 times. The 2AFC task will be administered online using PsychoPy (Peirce et al., 2019) and participants will be recruited using the Prolific.ac online experimental platform. The task takes approximately 5 minutes to complete.

The experimental items consisted of sentences in Spanish…

We piloted the 2AFC experiment on 100 native Spanish speakers to assure critical item quality and assess the difficulty of the task. Crucially, our existing data does not include our population of interest, L2 learners.

**LexTALE**. The LexTALE task (Lemhöfer & Broersma, 2012) is a lexical decision experiment in which the participants are presented real and nonsense words. We will administer the Spanish version, LexTALE-ESP (Izura, Cuetos, & Brysbaert, 2014). Participants have to decide if the word they are presented is real or not using the keyboard (‘1’ for real, ‘0’ for fake). The purpose of the LexTALE task is to get an assessment of the participants’ vocabulary size in Spanish, which is assumed to be a proxy for Spanish proficiency. The LexTALE task takes approximately 5 minutes to complete.

We administered the Lexical Test for Advanced Learners of Spanish (Izura et al., 2014; LexTALE-ESP, Lemhöfer & Broersma, 2012) in order to provide a standardized assessment of the participants’ proficiency/vocabulary size in Spanish. On this measure, scores can range from −20 to 60, with native speaker values generally above 50. Individuals with little or no knowledge of Spanish typically score from −20 to 0. For the purposes of the present study, proficiency is treated as a continuous variable, thus we consider a monolingual English speaker to have little to no proficiency in Spanish.

**Empathy Questionnaire**. Finally, participants will complete the Empathy Quotient (EQ, Baron-Cohen & Wheelwright, 2004), a 60-item questionnaire that presents likert-type items ranging from ‘strongly agree’ to ‘strongly disagree’. The EQ is scored to produce a single value indicating an individual’s level of empathy. The EQ takes approximately 5 minutes to complete.

## Procedure

We excluded data according to the following criteria: error during data collection; clear lack of understanding or engagement during the task, i.e., all ‘1’ responses, failed three attention checks, etc.; participants reporting having learned Spanish before the age of 13; participants with knowledge of language other and English and Spanish. Data from a total of 78 participants was discarded because the session timed our and/or data was incomplete. An additional 8 participants were discarded due to low accuracy (n = 5), incomplete data (n = 2), and failed attention checks (n = 1). A total of 224 participants met the criteria for inclusion.

## Statistical analyses

We report two primary statistical analyses that were pre-registered prior to collecting the learner data. First, we analyzed response accuracy using Bayesian hierarchical logistic regression. The model considered response accuracy for the population effects *utterance type* (declarative broad focus, declarative narrow focus, interrogative yes/no, interrogative -wh), proficiency (i.e., LexTALE score), empathy quotient, and the proficiency by empathy quotient interaction. The likelihood of the model was Bernoulli distributed with a logit link function. The criterion, *response*, was coded as “1” for correct responses and “0” for incorrect responses. Thus, the first analysis modeled the probability of responding correctly to the prompt “Is this a question”. We specified group-level effects for participants, speaker variety, and items.  
The slope for *utterance type* varied for the participant effect, as did the proficiency by empathy quotient interaction for the speaker variety effect. All continuous variables were standardized and ‘interrogative yes/no’ was set as the baseline for *utterance type*, thus the model intercept represents the probability of a learner with average proficiency and average empathy responding correctly to a yes/no question.

The model included regularizing, weakly informative priors (Gelman, Simpson, & Betancourt, 2017).

We established a region of practical equivalence (ROPE) of ± 0.05 around a point null value of 0 (see Kruschke, 2018). We will sample the posterior distribution of the model for statistical inferences. A region of practical equivalence (ROPE) will be established around a point null value of 0 (Kruschke, 2018) using the following formula:

For all models, median posterior point estimates are reported for each parameter of interest, along with the 95% highest density interval (HDI), the percent of the region of the HDI contained within the ROPE, and the maximum probability of effect (MPE). For statistical inferences, a posterior distribution for a parameter β in which 95% of the HDI falls outside the ROPE and a high MPE (i.e., values close to 1) were taken as compelling evidence for a given effect. All exploratory analyses, explicitly described as such, include posterior point estimates, the 95% HDI and the MPE. The analyses were conducted in R and models fit using the probabilistic programming language stan via the R package brms (Bürkner, 2017, 2018).

# References

Astruc, L., Vanrell, M. del M., & Prieto, P. (2016). Cost of the action and social distance affect the selection of question intonation in Catalan. In M. E. Armstrong, N. Henriksen, & M. del M. Vanrell (Eds.), *Intonational grammar in Ibero-Romance: Approaches across linguistic subfields* (pp. 93–113). John Benjamins Publishing Company. <https://doi.org/10.1075/ihll.6>

Baron-Cohen, S. (2011). *Zero degree of empathy. On empathy and the origins of cruelty*. London, England: Penguin.

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*(2), 163–175. <https://doi.org/10.1023/B:JADD.0000022607.19833.00>

Bishop, J., & Kuo, G. (2016). Do "autistic-like"" personality traits predict prosody perception? *Talk presented at LabPhon15 satellite workshop on personality in speech perception and production, ithaca, NY.*

Brandl, A., González, C., & Bustin, A. (2020). The development of intonation in L2 Spanish: A perceptual study. In A. Morales-Front, M. J. Ferreira, R. P. Leow, & C. Sanz (Eds.), *Hispanic linguistics: Current issues and new directions* (pp. 12–31). John Benjamins Publishing Company. <https://doi.org/10.1075/ihll.26>

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. <https://doi.org/10.18637/jss.v080.i01>

Bürkner, P.-C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*, *10*(1), 395–411. <https://doi.org/10.32614/RJ-2018-017>

Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology*, *96*(4), 671–681. <https://doi.org/10.1037/0022-0663.96.4.671>

Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and Brain Sciences*, *32*(2), 121–138. <https://doi.org/10.1017/S0140525X09000545>

Degen, J., & Tanenhaus, M. K. (2016). Availability of alternatives and the processing of scalar implicatures: A visual world eye-tracking study. *Cognitive Science*, *40*(1), 172–201. <https://doi.org/10.1111/cogs.12227>

Esteve-Gibert, N., Portes, C., Schafer, A., Hemforth, B., & D’Imperio, M. (2016). *The role of individual empathic skills on the online processing of intonational meaning*. Bilbao, Spain: Basque Center on Cognition, Brain; Language. <https://doi.org/10.13140/RG.2.2.19401.13926>

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*, 1–15. <https://doi.org/10.3758/s13421-019-00990-w>

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 459–473. <https://doi.org/10.1098/rstb.2002.1218>

Gelman, A., Simpson, D., & Betancourt, M. (2017). The prior can often only be understood in the context of the likelihood. *Entropy*, *19*(10), 1–13. <https://doi.org/10.3390/e19100555>

Henriksen, N., Armstrong, M. E., & Garcı́a-Amaya, L. (2016). The intonational meaning of polar questions in Manchego Spanish spontaneous speech. In M. E. Armstrong, N. Henriksen, & M. del M. Vanrell (Eds.), *Intonational grammar in Ibero-Romance: Approaches across linguistic subfields* (pp. 181–205). John Benjamins Publishing Company. <https://doi.org/10.1075/ihll.6>

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. <https://doi.org/10.1037/t47086-000>

Kruschke, J. K. (2018). Rejecting or accepting parameter values in bayesian estimation. *Advances in Methods and Practices in Psychological Science*, *1*(2), 270–280.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, *44*(2), 325–343. <https://doi.org/10.3758/s13428-011-0146-0>

Nieuwland, M. S., Ditman, T., & Kuperberg, G. R. (2010). On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities. *Journal of Memory and Language*, *63*(3), 324–346. <https://doi.org/10.1016/j.jml.2010.06.005>

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). <https://doi.org/10.5334/labphon.238>

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., … Lindelv, J. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203.

Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, *64*(4), 878–912. <https://doi.org/10.1111/lang.12079>

Portes, C., Beyssade, C., Michelas, A., Marandin, J.-M., & Champagne-Lavau, M. (2014). The dialogical dimension of intonational meaning: Evidence from French. *Journal of Pragmatics*, *74*, 15–29. <https://doi.org/10.1016/j.pragma.2014.08.013>

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

Trimble, J. C. (2013a). *Acquiring variable L2 spanish intonation in a study abroad context* (PhD thesis). University of Minnesota.

Trimble, J. C. (2013b). Perceiving intonational cues in a foreign language: Perception of sentence type in two dialects of Spanish. In C. Howe (Ed.), *Selected proceedings of the 15th hispanic linguistics symposium* (pp. 78–92). Somerville, MA: Cascadilla Proceedings Project.

Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, 1–30. <https://doi.org/10.1017/S0272263106060013>