JOURNAL ARTICLE

Individual empathy levels affect gradual intonation-meaning mapping: The case of biased questions in Salerno Italian

Riccardo Orrico[1,2] and Mariapaola D'Imperio[3]

[1] Department of Humanities, University of Salerno, IT

[2] Aix Marseille University, CNRS, Aix-en-Provence, FR

[3] Department of Linguistics and Cognitive Science Center (RuCCS), New Brunswick, NJ, US

Corresponding author: Riccardo Orrico (rorrico@unisa.it)

The paper investigates the interplay between intonational cues and individual variability in the perceptual assessment of speakers' epistemic bias in Salerno Italian yes-no questions. We present a perception experiment in which we manipulated pitch span within the nuclear configuration (both nuclear accent and boundary tone) to predict degree of perceived positive bias (i.e., expected positive answer) to yes-no question stimuli. Our results show that a wider pitch span within the nuclear region predicts a higher degree of perceived positive bias, while negative bias is predicted by narrow pitch span. Crucially, though, two interacting sources of listener variability were uncovered, i.e., prolonged exposure to a non-native dialect as well as degree of empathy (i.e., Empathy Quotient, EQ). Exposure to non-native phonological systems was found to affect the way pitch span is mapped onto perceived epistemic bias, through category interference, though mediated by EQ levels. Specifically, high-empathy listeners were more affected by degree of non-native dialect exposure. EQ scores were hence found to have an effect on gradual span manipulation by interacting with the dialect exposure effect. These results advance our understanding of the intonation-meaning mapping by taking into account both the impact of gradual phonetic cues on meaning processing as well as uncovering sources of cognitive variability at the perceiver's level.

1. Introduction

Linguistic research within the last decades has uncovered the importance of intonation in spoken communication, and specifically its role in shaping pragmatic meaning, such as conversational implicatures. In a seminal study, Ward and Hirschberg (1985) proposed that a specific rise-fall-rise contour in American English can implicate uncertainty, as in the example 'He's a good badminton player' shown in **Figure 1**. Note that pitch rises on 'badminton' and then again later, after an intermediate fall, at the end of 'player.' What is implied here is that the subject is not a really good player, with the consequence that the propositional content might be analyzed as being not true by the interlocutor.

An early variant of the rise-fall-rise contour depicted in **Figure 1**, i.e., the classic 'contradiction contour,' appears to be used to reject the truth value of a proposition, as in the classic example mentioned by Liberman and Sag (1974), reported in (1b) below with a modern ToBI-style notation (Beckman & Ayers, 1997). Notice in B's reaction to A's
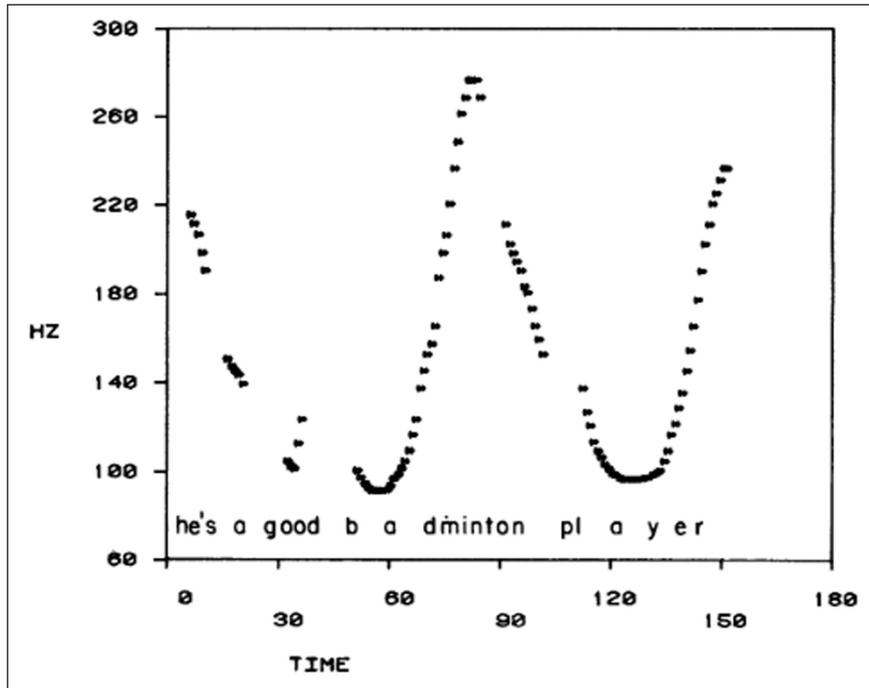
Figure 1: Example of the rise-fall-rise uncertainty contour taken from Ward and Hirschberg (1985).

statement that the pitch falls onto the first stressed syllable of the utterance, i.e., [ˈtaɪ] of 'elephantiasis' followed by a low plateau and a final low rise starting from the last stressed syllable, i.e., [ˈkyʊ] of 'incurable.'

(1)   A*:*   My fate is sealed. I've been diagnosed with elephantiasis

         H + !H*   L* L-H%

   B:
         Elephantiasis isn't incurable.

In later work, Pierrehumbert and Hirschberg (1990) have argued that this rise-fall-rise contour is not necessarily confined to contradictory meanings. In fact, it appears that this version of the rise-fall-rise can also be found in tag questions as in (2), in which case the contour brings an element of uncertainty (as opposed to having a fall), as observed by Cruttenden (2008), though this author uses a British-style prosodic annotation (not reported here).

(2)   He's passed hasn't he

Specifically, superimposing what Cruttenden defines as a final 'low rise' (which is the nuclear part of the contour) on the tag, as opposed to a fall, might be perceived as if the speaker is uncertain about the proposition evoked by the question (i.e., that the subject has actually passed). Note that Ward and Hirschberg (1986) claimed that this type of rise-fall-rise can be assigned the broad meaning of 'lack of speaker commitment to a scale,' which can convey either speaker uncertainty or speaker incredulity. Later, Hirschberg and Ward (1992) reported that the two sub-meanings of rise-fall-rises can be discriminated in perception on the basis of pitch range information, with a wider range being associated with incredulity.

Note that the role of pitch range, and specifically pitch span, i.e., the f0 excursion size within either pitch accent or edge tone, has traditionally been taken to be paralinguistic (Liberman & Pierrehumbert, 1984; Pierrehumbert & Beckman, 1988) given that it can gradually vary due to factors such as physical distance between the interlocutors or emotional factors. It is only recently that pitch scaling has been found to affect categorical perception of modality (cf. Vanrell, 2011), in that it can distinguish questions from statements within rise-fall pitch accents sharing the same alignment values. In this study we investigate two open issues. One is the relationship between gradualness in the phonetic input and gradualness in the intonational meaning mapping. The other is how generalizable this mapping is, given that individual listeners are endowed with a specific set of cognitive skills. Specifically, here we argue that a certain type of pragmatic meaning, i.e., epistemic bias (expectation about a positive or a negative answer to a polar question), can be gradually signaled and inferred through pitch span modifications. Crucially, here we show that the direction of the effect depends not only on the tonal composition of the pitch accent and long-term exposure to non-native dialects, but also on a specific individual cognitive skill, i.e., degree of empathy. In other words, on one side, gradualness of phonetic cue usage is argued to be directly tied to gradualness in the expression of pragmatic meaning. On the other, we show that the picture is rendered even more complex by the fact that, from a processing point of view, gradualness on both sides of the intonation-meaning mapping cannot be modeled in a uniform fashion within a linguistic community, given a strong interaction between long-term exposure effects and empathy levels on the part of the perceiver.

Previous findings about the use of pitch span in polar questions were already reported in Savino and Grice (2011) for Bari Italian, in which a wider excursion within the nuclear pitch accent was found to be associated with negative bias (i.e., expectation of a negative answer). Similar findings for Catalan reported by Borràs-Comes, Vanrell, and Prieto (2014) showed that pitch excursion is the main cue to differentiate a counter-expectational question (wider excursion) from a statement (narrower excursion), despite the fact that the two speech acts share the same rise-fall nuclear contour. Moreover, in line with usage-based models, exposure to different dialects and languages can impact this mapping. Note that this goes counter to recent universalist views of the use of gradual use of pitch height and span in signaling uncertainty and attitudinal meaning (Gussenhoven, 2002).

The role of intonation in signaling epistemic bias, among the other contributions, has been recently compared to the function of lexical epistemic operators in a corpus study on declarative questions in English and French by Safarovà (2006) (see also Section 3 below). Gunlogson (2008) has later adopted the notion of *contingent commitment* on the part of the addressee to account for the restricted distribution of rising declaratives, so that bias is explained by variable commitment to the truth of a proposition. In the present study we will concentrate on intonationally-induced bias in yes-no questions. Here we employ Ladusaw's (2004) definition of biased question as a question in which the speaker is predisposed to accept one particular answer as the right one. Moreover, we use the term positive bias to indicate that the speaker is biased towards a positive answer (a 'yes' answer) while a negative bias would signal an expectation for a negative answer (a 'no' answer).

Though speakers' epistemic disposition towards a proposition, such as the degree of certainty towards its content, can also be expressed through lexical marking (for instance, through adverbials such as 'surely,' 'certainly'), intonational form has been recently found to be a relevant cue (Vanrell, Mascaro, Torres-Tamarit, & Prieto, 2013 for confirmation-seeking questions in Catalan), as well as facial and body gestures (Palmer, 2001; Borràs-Comes, Kiagia, & Prieto, 2019). Confirmation-seeking questions, for instance, can be defined as questions for which the speaker shows a bias which can be calculated on the

basis of private beliefs combined with world knowledge and information made available by the discourse background (cf. Bolinger, 1989; Büring & Gunlogson, 2000). Note also that question-answer pairs have also been studied in an interactional perspective that capitalizes on spontaneous or semi-spontaneous speech (see for instance Enfield, Stivers, & Levinson, 2010).

Crucially though, the suggested link between intonation and meaning postulated in the aforementioned studies is taken to be community-homogeneous as well as categorically expressed by the use of a specific phonological tune. In other words, according to these studies, given a certain syntax-prosody combination, each listener within a specific linguistic community would process the same form-meaning mapping. But prosody and intonation are part of the phonological grammar of a talker and we know that speech encodes different kinds of socio-indexical information, though this area has been mainly investigated within segmental phonetics and phonology (Johnson, Strand, & D'Imperio, 1999, Perry, Ohde, & Ashmead, 2001; Pierrehumbert, Bent, Munson, Bradlow, & Bailey, 2004; Drager, Hay, & Walker, 2010, inter alia). The role of social and indexical properties in intonation variation has only received very limited attention, and mainly from the point of view of production (see Clopper & Smiljanic, 2011 and Levon, 2016 for gender effects).

Another area that has received little attention in intonation studies is the role of long and short-term exposure to a non-native dialect. Individual phonological systems can vary depending on the amount of non-native dialect exposure, as predicted both by exemplar models (e.g., Goldinger, 1998; Johnson, 1997, 2006; Pierrehumbert, 2001, 2002) and phonological theories on Second Language Acquisition (Flege, 1995; Best & Tyler, 2007; Mennen, 2015). A recent study by Orrico, Cataldo, and D'Imperio (to appear), discusses the effect of dialect exposure in the production of question tunes in Salerno Italian, which was found to mainly affect the realization of H% versus L% boundaries in yes-no and wh-questions. Additionally, it has been very recently shown that individual cognitive variability has an effect on the online processing of intonational meaning, specifically through the effect of pragmatic skills (Estève-Gibert et al., 2020) measured through the Empathy Quotient (see Section 1.4).

Note that in Italian there are no morpho-syntactic means to formulate a yes-no question. While morpho-syntactic strategies are available to differentiate an imperative from other modalities, interrogatives and declaratives are typically realized with the same SVO syntactic structure (Sobrero, 1993; Rossi, 2015) so that intonation and prosody are the only means to recover illocutionary meaning (cf. D'Imperio 2000; 2002; Petrone, 2008; see Cangemi & D'Imperio, 2015 for speech tempo effects). Hence, the present investigation seeks evidence for the role of intonation in perceiving speakers' bias in yes-no question tunes of Salerno Italian, by specifically addressing quantitative pitch span effects on one side and individual perceiver differences on the other. The novelty of the present study is that we simultaneously investigated the combined effect of exposure and empathy on the gradual pitch span manipulation to better understand the impact of individual variability on the intonation-meaning mapping. Also, by employing resynthesized stimuli, we set out to explore fine-grained phonetic detail on phonological encoding of rising intonation in a controlled way. More specifically, we test the hypothesis that the effect of pitch span on epistemic bias response is modulated by the interaction between individual empathy level and non-native dialect exposure. This work thus represents an important step towards a fuller integration of intonation and pragmatic theories by including talker/perceiver variability which we claim to be relevant in determining tune-meaning association.

In order to appreciate how both phonological and phonetic intonation cues affect the interpretation of epistemic stance in question intonation and how our results interact with the literature, we need to understand the issues involved in the analysis and current theories of intonational meaning. We therefore begin by summarizing contemporary theories of intonational meaning in Section 1.1. Then, in Section 1.2, we discuss the relevant background on the issue of epistemic stance and intonation, which, as we will see, plays a major role in responses to polar questions. In Section 1.3 we address the issue of gradual versus categorical cues to intonational meaning followed by a section on the relevance of individual variability in speech perception, with a special look to prosody and intonation. Having introduced the relevant theoretical background, we then briefly introduce the specificities of Salerno Italian intonational phonology in Section 2 and then move on, in Sections 4 and 5, to the details of our perceptual experiment. Finally, in Section 6 we provide a discussion of our results in terms of current models of phonological representation and processing.

1.1. Intonational meaning and the Autosegmental Metrical Theory

At the heart of the Autosegmental Metrical (AM) model of intonation (Pierrehumbert, 1980; see Ladd, 2008 for a review), there is the assumption that an intonation contour is a sequence of discrete tonal elements, i.e., tone targets, combining into either monotonal or bitonal pitch accents and edge tones (phrase accents and boundary tones). These elements originate from two discrete primitive level tones, high (H) and low tones (L). Studies conducted in the last 40 years within this framework have allowed the linguistic community to reach a good understanding of the phonology-phonetic interface at the prosodic level. Nevertheless, several controversies still remain unresolved, especially as far as the specific role of intonation in the expression of pragmatic meaning, what is referred to as 'intonational meaning' by Ladd (2008).

Probably the best known intonational meaning model is the one proposed by Pierrehumbert and Hirschberg (1990), in which the authors propose a strong tune compositionality. In this model, each single tone recoverable from the f0 contour would carry an inherent, context-independent meaning. In line with the tenets of AM, the model assigns specific meanings to the two primitive tones, H and L, and their contribution in pitch accent and edge tone combinations. The pragmatic meaning of an utterance is then obtained by adding the meanings of each tone composing the tune. Additionally, the authors posit a hierarchical phonological model in which the three structural elements composing a tune—pitch accents, phrase accents, and boundary tones—operate over a progressively higher domain dictated by the node they are attached to. Therefore, pitch accents would have scope over the lexical item they are attached to, phrase accents over the intermediate phrase, and boundary tones over the entire intonation phrase.

The strong compositionality of existing AM-based models shows nevertheless some serious drawbacks. Positing that each single tone possesses an inherent meaning makes it difficult to account for collocations of nuclear pitch accents and boundary tones (see Dainora, 2001). That is why several recent models of intonational meaning, couched within AM, have put forth a weak compositional approach, as Steedman's (2007) model for English intonation and Portes and Beyssade's (2015) account for French. We believe that both compositionality and the notion of 'inherent' meaning of an intonational contour need to be revised. On one side, intonational meaning can be viewed as gradual and not simply as discrete. Similar to the phonetics-phonology mapping implied in exemplar theory (Goldinger, 1998; Johnson 1997; 2006; Pierrehumbert 2001; 2002), pragmatic meaning can be thought of as being gradually implemented both in production and perception.

This view would differ from a paradigm such as the one invoked by the Biological Codes theory proposed by Gussenhoven (2002), in which gradual acoustic cues would merely pertain to universal meanings, while linguistic intonation representations would only be discrete in nature (see Section 1.3 below).

1.2. The epistemic nature of intonational meaning

Very recent models argue that intonation is concerned with meaning related to the dialogical status of an utterance, such as its role in adding information to the Common Ground, its relation to private beliefs and the degree of commitment to a proposition, be it stated or evoked in the utterance (see Prieto, 2015, for a detailed account). A series of intonational meaning studies has hence connected specific intonational features to the role of carrying information relative to epistemic commitment on the part of the interlocutors. In some of these studies, the speaker's use of intonation is assumed to carry the function of expressing that she is taking responsibility for the truth of a proposition and/or she is attributing such responsibility to the addressee (Beyssade & Marandin, 2006; Krifka, 2017). Among these studies, Bartels (1999) proposes that the use of the L- phrase accent in American English signals commitment on the part of the speaker relative to the truth of a relevant proposition in the discourse and, at the same time, to an instruction aimed at the addressee to publicly committing to it. A different perspective, though within a similar framework, has been taken by Steedman (2007) for English and, more recently, by Portes and Beyssade (2015) for French. Both of these models posit, for instance, that the role of boundary tones in declarative sentences is to variably attribute commitment to either the speaker (when falling) or the addressee (when rising).

Additionally, some authors also propose that pitch accents are responsible for determining whether or not the information contained in the utterance is predicted by the speaker to be either consensual or controversial. In this view, intonation is taken to convey an important role in dynamically changing and managing the Common Ground, and this by adding, accepting, or rejecting a proposition (Krifka, 2017). Some work has also uncovered a complex interaction between syntactic markers and illocutionary force (i.e., declarative versus interrogative modality) due to the presence of a specific intonational tune. For instance, in a study of declarative questions in American English, Gunlogson (2004) claims that rising nuclear contours in declaratives may indicate lack of commitment to the proposition on the part of the speaker, while rises in interrogative sentences might attribute commitment to the addressee. In this view, the meaning of declarative questions would be derived compositionally and in light of the direction of the terminal part of the contour. Specifically, a terminally falling declarative (statement tune) would signal speakers' commitment to the proposition while a rising contour (yes-no question tune) would signal lack of the same (hence commitment being attributed to the addressee).

A great deal of research operating within the intonation-meaning framework mentioned above has also investigated the role of intonation in expressing epistemic bias in questions. Among various kinds of pragmatic bias, epistemic bias (Sudo, 2013) pertains to presupposed information about the content of the answer to the question on the part of the speaker. Within different speech acts, biased questions belong to the set of non-canonical questions, in that their function is not just a way to elicit information, but also to convey an epistemic bias (i.e., an expectation) about what that information might be (Dayal, 2016). While, according to standard semantics, a canonical yes-no question creates a balanced partition between the two polarities of a proposition, the role of epistemic bias is that of creating an unbalanced partition, with the speaker leaning towards either p (positive bias) or ¬p (negative bias). Several models have accounted for the role of bias as attributing question properties that are typical of assertions. Asher and Reese (2007), for

example, analyze biased questions as complex speech acts, formed by both an assertion and a question part. A typical example of biased question in their analysis is represented by tag questions, which overtly include both a declarative and an interrogative component. Apart from tag questions (also treated by Ladd, 1981), bias appears to be an intrinsic property of negative polar questions (Ladd, 1981; Büring & Gunlogson, 2000; Asher & Reese, 2005; 2007; see also Goodhue & Wagner, 2018) and declarative questions.

From this short review, note that the role that intonation plays in determining the presence and the polarity of the bias (positive or negative) has emerged to the frontline. Analyses like the one proposed by Asher and Reese (2005; 2007) argue that the presence of an epistemic bias can be signaled by the use of intonational cues alone. Similar theoretical claims have been tested in recent experimental studies. Among them, Michelas, Portes, and Champagne-Lavau (2016) designed a production study aimed at investigating the role of intonation in encoding speaker commitment in French. Speakers were asked to produce questions in two pragmatic conditions, i.e., unbiased and negatively biased questions. Findings show that the role of pitch accents is crucial in differentiating across the two conditions, with H* nuclear accents signaling unbiased questions and H+!H* accents signaling negative bias, showing the contribution of intonation in determining the beliefs state of speakers. Portes, Beyssade, Michelas, Marandin, and Champagne-Lavau (2014) also investigated the dialogical function of intonation in French by specifically addressing the issue of how intonational contours express speaker commitment and 'call on addressee,' which is a way to anticipate the listener's reaction to a speaker's utterance. In particular, the authors found that intonational contours elicit specific reactions from the addressee, showing the role of intonation in determining whether information is agreed or not among conversational participants, and hence added to the Common Ground.

Further support for the role of intonation in conveying information regarding commitment and agreement between interlocutors has been provided by empirical work on other Romance varieties, such as Puerto Rican Spanish (Armstrong, 2012) and Catalan (Prieto & Borràs-Comes, 2018). In these studies, the authors show that question intonation can encode fine-grained distinctions related to the speaker's own commitment to the proposition expressed through some form of epistemic stance about the conveyed content (Heritage, 2012). Prieto and Borràs-Comes (2018) designed a perception task in which native speakers of Catalan were asked to rate the acceptability of tune-context pairs. Four yes-no question contours were analyzed, a L* H% rise, a H+L* L% fall, a L+¡H* L% rise-fall, and a L+H* LH% rise-fall-rise, which were paired with six different contexts, resulting in different degrees of speaker commitment and agreement. Results show that intonation does indeed function as a device expressing commitment and agreement. Specifically, they show that the falling H+L* L% contour was rated as highly acceptable in high commitment contexts, while the L* H% rise correlates with low levels of agreement.

Intonational studies on both Catalan and Bari Italian have also pointed out the role of pitch direction (rises versus falls) in signaling different types of questions, such as the difference between confirmation and information-seeking questions. Grice and Savino (1997), for instance, showed that Bari Italian yes-no questions (questions about new information) are marked by the presence of a nuclear rising L+H* accent while confirmation-seeking questions (in which the information addressed is given) are generally marked by the presence of a falling nuclear H+L* accent. Hence pitch accent category and/or edge tone combinations appear to have a major role in signaling bias in questions.

1.3. Gradient intonational cues to meaning

One important issue concerning the nature of intonational meaning is the divide that is often argued to exist between gradient versus categorical elements of intonation. Within traditional AM approaches (Liberman & Pierrehumbert, 1984), pitch range modulations

within the two phonological, L and H, primitive tones are argued to be used only to express paralinguistic meanings (such as emphasis, degree of speaker involvement, etc.). As a consequence, according to this view, a phonological tune can be expressed with very different pitch ranges without changing its pragmatic function. Moreover, pitch range can vary along two dimensions (see Ladd, 2008 for a discussion), i.e., span, which is the excursion between f0 peaks and valleys, and level, which relates to the register (or key) within which target f0 values are realized. Variations within these two components of intonation have been linked to the expression of paralinguistic meanings such as arousal, involvement of the speaker, and social dominance (see Ladd, 2014). Nevertheless, the fact that gradient cues associate with paralinguistic meanings and intonational categories with linguistic meaning is quite controversial. As pointed out by both Grice and Bauman (2007) and Prieto (2015), the separation between the two dimensions of cues and types of meanings is far from clear-cut.

A theory which attempts to address both the paralinguistic meaning and the grammatical meaning of intonation is the Biological Codes theory, proposed by Gussenhoven (2002). The theory is built on the assumption that the language-specificity nature of intonational meaning is the result of a process of grammaticalization of universally-valid phonetic-meaning mappings. Building on Ohala's (1983) ethological view of fundamental frequency use (Frequency Code), Gussenhoven extends this view by postulating four Biological Codes from which two types of meanings are generated, one that is affective—signaling attributes of the speaker—and an informational, linguistic, meaning, coding information related to the message. In line with Ohala, the Frequency Code is intended as directly relating pitch height to the size of a speaker's vocal folds, giving rise to meanings such as submissiveness and politeness (high pitch), versus impoliteness and confidence (low pitch). The informational interpretation of the Frequency Code relates to the opposition between finality and non-finality and certainty versus uncertainty, with the second attribute in both cases being associated with higher pitch. By the same token, assertiveness versus questioning would hence be related to f0 rises or increased peak height (see for instance Chen, Gussenhoven, & Rietveld, 2004). Note though that the Frequency Code predictions for politeness, being related to rising or high pitch, have recently been challenged by empirical results in a variety of languages (see Chikulaeva & D'Imperio, 2018 for Russian; Winter & Grawunder, 2012; and Brown, Winter, Idemaru, & Grawunder, 2014 for Korean).

The link between pitch span and linguistic, pragmatic meaning has already been highlighted by several studies in the past. As mentioned above, Hirschberg and Ward (1992) were the first to point out that that information relative to pitch span is actually instrumental for determining the difference between an incredulous or an uncertain reading of the rise-fall-rise contour in American English. Specifically, having ruled out the hypothesis of polysemy of this contour, they proposed that the two readings might represent a subset of the wider meaning labeled as 'lack of speaker commitment.' Moreover, recent evidence from Romance languages points to the fact that pitch span can indeed express linguistically structured meanings. Borràs-Comes et al. (2014), for instance, show that pitch span variation within the rising L + H* pitch accent can encode three different meanings in Catalan, i.e., neutral statement (lower peak), narrow-contrastive focus (medial peak), and counter-expectational yes-no question (higher peak).

Similarly, evidence for several varieties of Italian (Gili Fivela et al., 2015) points to the role of wider pitch span in the nuclear configuration region of wh-questions being used to encode counter-expectational meaning. The same type of phonetic cue appears to be employed in the Italian variety spoken in Este, Padua (a Northern variety) to signal incredulous echo wh-questions (Crocco & Badan, 2016). In Salerno Italian, which is the variety employed in the present study, this specific meaning appears to be brought about

by the combination of a nuclear rising L* + H pitch accent with expanded pitch excursion (Orrico & D'Imperio, 2020). Interestingly, the meanings which are cross-linguistically attributed to expanded pitch span—i.e., incredulity, counter-expectation, negative bias— all imply some sort of rejection of the propositional content conveyed by the utterance.

Additionally, and more importantly, in all these cases it might be possible to talk about gradience in Bolinger's (1989) sense, which implies a mapping between gradient phonetic cues and gradient meanings. Indeed, apart from paralinguistic meanings, also linguistic ones might be conceived in a gradient way. Heritage (2012), for example, argues that the same propositional content can be modulated by a language grammar on the basis of the relative knowledge possessed by the speaker and the addressee. This would lead to a continuum between unbiased yes-no questions, representing that the speaker is not committing to any proposition, and statements, representing an increased knowledge and commitment on the part of the speaker. In this view, biased questions would occupy an intermediate position within the continuum.

1.4. Individual variability in the tune-meaning mapping

It is well established that speech encodes indexical information about the speaker producing it, and this appears to be true also for intonation. In a study by Clopper and Smiljanic (2011), effects of both gender and regional dialect (Southern versus Midland American English) on intonation production were found. Specifically, the authors reported that female speakers of both dialects produce more L* + H accents (while males show a preference for H* in the same conditions). Crucially, though, they also found an inter-dialect difference in the frequency use of L- phrase accents for female speakers only, showing that the way different social variables interact with each other is a relevant aspect. An effect of gender has also been found in the production of high-rise terminals in London English by Levon (2016), a dialect in which men show a tendency to use high-rises in narratives to draw attention to interesting or new elements in the discourse. On the other hand, women appear to use this contour to maintain epistemic authority over the narrative.

Studies on intonational variability have unveiled that group- and speaker-specific strategies are also employed at the level of intonational encoding of pragmatic meanings. What is more, some of these studies have also highlighted the role of variability at the listener level. For example, Arvaniti, Baltazani, and Gryllia (2014) found that the pragmatic meaning of wh-question tunes in Greek is interpreted differently not only as a result of speaker gender, but also depending on the gender of the listener, with an increased tendency towards interpreting a wh-question as a statement with negative implicatures if the speaker is male and the listener is female. Additionally, work by Cangemi, Krüger, and Grice (2015) shows that different speakers use different prosodic strategies to encode three different types of focus in German (broad, narrow, and corrective) and, additionally, different listeners also use different strategies to encode them (though they do not elaborate on possible sources of such speaker- and listener-specific differences).

These studies point to the fact that both listener and speaker idiosyncrasies can affect the complex form-meaning mapping of intonation, contrary to what is traditionally maintained (e.g., Pierrehumbert & Hirschberg, 1990). Apart from indexical information, the body of studies dedicated to individual variability has also shown that two other factors appear to affect the linguistic behavior of speaker and listener, i.e., social variability and individual cognitive traits (Kidd, Donnelly, & Christiansen, 2018; Yu & Zellou, 2019). Exposure to different dialects, in terms of type and amount, is another important individual variable. Chen, Rattanasone, Cox, and Demuth's (2017) eye-tracking study, for example, shows that Australian English listeners with and without early exposure to other English dialects

show different behavior in reaction to mispronounced vowels. Specifically, listeners with exposure to other varieties appear to be more tolerant to deviant exemplars, which leads the authors to conclude that phonological categories in these individuals might be less rigid than those for individuals without exposure. More recently, Levy and Hanulíková (2019) report that variability in vowel production on the part of children is related to their levels of non-native dialectal exposure.

Only scarce evidence of the effect of exposure to other phonological systems can be found in intonation studies. In a study on tonal alignment acquisition in L2, Mennen (2004) reported that adult Dutch learners of Greek shift their tonal alignment patterns in prenuclear rising pitch accent towards that of native speakers of Greek, though not enough to obtain Greek-like values. Interestingly, the same learners would present alignment values in their L1 that would not be native-like either, pointing to a bi-directional effect of intonational convergence. Note that Gili Fivela et al. (2015) put forward the hypothesis that contact among regional dialects of Italian might be an important source for the amount of variability in intonational patterns. In a study on tonal alignment imitation, D'Imperio, Cavone, and Petrone (2014) report that short-term exposure to a different alignment in a question rise-fall contour produced by a speaker of a different variety of Italian can be immediately stored and reproduced.

Very recently, in a study on the perception of Corsican French intonation, Portes and German (2019) found that exposure to Continental French modulates the response to a regional pattern cueing either a question or a statement. This echoes findings by Walker, Szakay, and Cox (2019) in which socio-indexical segmental cues to Australian versus New Zealand English appear to be modulated by the amount of exposure to each variety. In fact, the authors find that the expected regional priming effect is actually weaker in their study than the one found in previous studies (cf. Hay & Drager, 2010). In a similar vein, Warren's (2017) findings on sociophonetic priming in the perception of uptalk in New Zealand English point to a mutual interaction of dialectal exposure and different socio-indexical sources of priming.

From this short review, it emerges that individual variability in intonation research has mainly concentrated on socio-indexical and regional features. Only very recently individual cognitive skills have been argued to play a role in tone and intonation perception variability, both in L1 and L2. Among those skills, musicianship has been found to have a major role in terms of L2 acquisition in both children and adults. In a series of ERP and perception studies, general musical training (musicianship) has been linked to faster novel word learning by French participants, both in terms of form and meaning (Dittinger et al., 2016; Dittinger, Valizadeh, Jäncke, Besson, & Elmer, 2018). Among musical skills, a recent study has shown that rhythmical aptitude predicts performance in correct stress placement in French learners of L2 English (Cason, Marmursztejn, D'Imperio, & Schön, 2019). In addition to musicianship, other individual cognitive traits appear to affect prosodic cue processing. Jun and Bishop (2015), for instance, showed that individuals with a high level of autistic traits (measured through the Autistic Quotient) appear to be more sensitive to prosodic boundaries in cueing high or low attachment of a relative clause, showing that such cognitive factors might have an effect on syntactic parsing. In addition, these results are in line with findings relating higher degree of pragmatic skills and better discrimination and exploitation of pitch accent contrast in meaning processing (Colon & Bishop, 2015; Bishop, 2016). Along these lines, Estève-Gibert et al. (2018; 2020), in an Eye Tracking experiment on contrastive meaning processing in French, found that individuals who scored lower at the Empathy Quotient test (Baron-Cohen & Wheelwright, 2004) rely primarily on lexical disambiguating cues instead of early intonational ones for the purposes of homophone disambiguation.

As we know, the covariation of phonetic cues and socio-indexical information is a hallmark of exemplar models of speech perception (Pierrehumbert, 2001; German, Carlson, & Pierrehumbert, 2013). In this study we will extend our view to cognitive and dialectal exposure effects to predict the influence of gradual pitch span variability on epistemic bias identification. But before we detail the experimental design of this study, we will briefly review the specificities of Salerno Italian intonational phonology.

2. Yes-no question intonation in Salerno Italian

Despite a great number of investigations having been dedicated to the intonational phonology of Italian regional varieties (see Grice, D'Imperio, Savino, & Avesani, 2005; Gili Fivela et al., 2015; D'Imperio, Baltazani, Gili Fivela, Post, & Vella, in press), quite a small number of studies have been dedicated to intonation in the variety spoken in Salerno (a Southern city, close to Naples). The only two studies that treat the intonational phonology of this variety within the Autosegmental-Metrical approach (Gili Fivela et al., 2015 and Orrico, Savy, & D'Imperio, 2019a) have unveiled high levels of variability in the use of yes-no question tunes. Specifically, analyzing data collected using the Discourse Completion Task (Blum-Kulka, House, & Kasper, 1989), Orrico et al. (2019a) showed that a yes-no question in Salerno Italian can be tonally expressed using a number of different configurations. Crucially, the investigations reported in their study showed that an $L+H^*$ pitch accent is the most frequent accent found in nuclear position (though L^*+H accents were also reported). The $L+H^*$ accent is realized as a rise starting in the pre-stressed syllable, with the peak aligned early in the stressed vowel and a subsequent fall which is completed within the same vowel. Additionally, both a falling-rising (HL-H%) and a falling (HL-L%) terminal can be employed, as shown in **Figure 2** below, though the $L+H^*$ HL-H% nuclear configuration is the most frequent one.

Orrico et al. (2019a) also provide an analysis of the distribution of these tunes according to four pragmatically different conditions, i.e., information-seeking, confirmation-seeking, echo, and counter-expectational questions. Crucially, no clear picture emerges when trying to map different contours onto different pragmatic contexts. Rather, as also suggested in Gili Fivela et al. (2015) for several varieties of Italian, there is a many-to-many mapping between tunes and contexts. Orrico et al. (2019a) hence propose that such high levels of variation depend on individual speaker variability. In other words, speakers might tend to use different subsets of tunes in specific contexts due to an individual preference for specific pitch accents and/or boundary tones across pragmatic categories. One additional result reported in the production study by Orrico et al. (2019a) concerns the distribution of tunes for counter-expectational questions, which are almost exclusively realized by means of a $L+H^*$ HL-L% tune. Additionally, while the same tune can be also used to express other pragmatic functions (i.e., unbiased yes-no questions as well as narrow focus
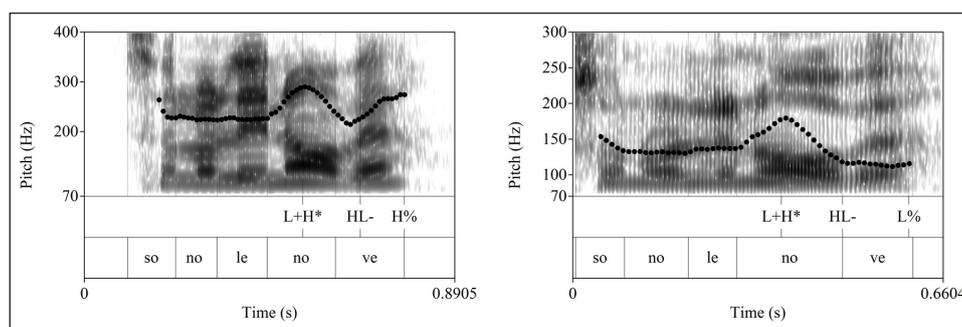

Figure 2: F0 contours for the question *Sono le nove?* ('Is it nine o' clock?') uttered with two different tunes (taken from Orrico et al., 2019a).

statements), it appears that differences in pitch span are crucial to discriminate a counter-expectational question from other types of uses of the same contour. **Figure 3** shows differences between the use of a L + H* HL-L% in counter-expectational questions and confirmation-seeking questions.

As shown in the pictures below, the incredulity meaning associated with a counter-expectational question appears to be conveyed through a narrower pitch span, as it can be noticed in the L + H* nuclear accent in the left panel utterance in **Figure 3**. This difference, which appears to reflect also a pragmatic difference, would not be captured by traditional AM models for intonational meaning, which posit that gradual span variations are not pragmatically relevant. The association between narrow span and counter-expectational meaning is quite interesting since it appears that in other Italian varieties, as well as in wh-questions in Salerno Italian, the incredulity meaning is on the other hand conveyed by a wider excursion of the pitch accent rise. Both in Italian and in other languages, an expanded pitch span has in fact been linked to meanings related to surprise, incredulity, and negative epistemic bias in questions on the part of the speaker, which would also be compatible with the predictions of Gussenhoven's (2002) Effort Code. Results for Salerno Italian yes-no questions, however, appear to go in the opposite direction, though Orrico et al. (2019a) only report a tendential behavior, so specific investigations should be carried out to verify this point.

3. Objectives and hypotheses

A perception experiment was hence designed to test the effect of variations in pitch span, in both nuclear pitch accent and boundary tone regions, on the perception of speaker bias in yes-no questions. In doing so, we evaluated the interaction with two individual variables, one social and the other linked to cognitive skills. Specifically, both long-term exposure (more than one year) to other phonological systems and listeners' Empathy Quotient (EQ) score (i.e., a cognitive skill related to pragmatic abilities) were tested. We will simply refer to these variables as, respectively, Exposure and EQ. The general hypothesis tested here (H1) is that pitch span can signal, in a gradual way, different types and degrees of speaker epistemic bias in a yes-no question. More specifically, we hypothesized, according to the tendency observed in the production results from Orrico et al. (2019a) and Orrico and D'Imperio (2020), that a narrower span would induce a perceived negative bias, i.e., the expectation of a negative answer to the question, while a wider span would induce a positive bias, i.e., the expectation of a positive answer, and that the effect would be the same for both pitch accent and boundary tone. Our secondary hypothesis (H2) was that both the social (Exposure) and the cognitive (EQ) variables would modulate the perceptual effect due to pitch span modification, though
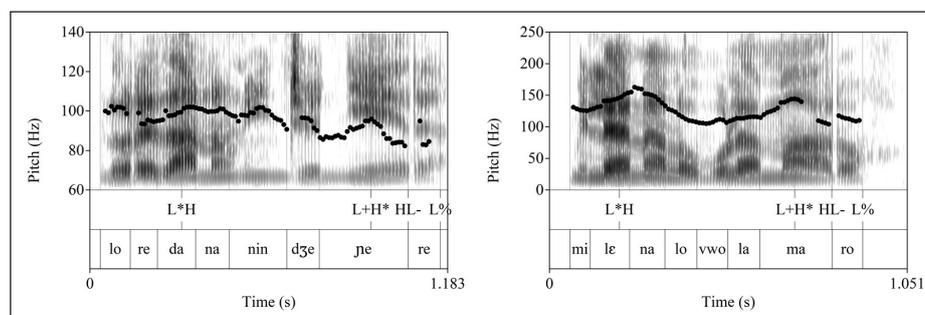

Figure 3: F0 contours for the counter-expectational question *Loredana un ingegnere?!?* ('Loredana (is) an engineer?!?') (left panel) and for the confirmation-seeking question *Milena lo vuole amaro?* ('Does Milena take it [the coffee] black?') (right panel) produced by the same male speaker.

in different ways. Specifically, while higher levels of Exposure would predict a higher response variability, higher EQ scores would predict a stronger attunement of the listener to the pitch cues employed to express epistemic bias. This last prediction is in line with evidence linking high EQ scores to higher levels of identification of other people's epistemic states (Lawrence, Shaw, Baker, Baron-Cohen, & David, 2004) as well as a more active use of pitch cues by high-empathy listeners in intonation meaning processing (Estève-Gibert et al., 2020).

4. Method
4.1. Stimuli
Perception stimuli were created by resynthesizing two yes-no questions uttered by a female native speaker of Salerno Italian (SI). The original set of stimuli from which the base stimuli were drawn was the same as the one used in Orrico, Savy, and D'Imperio (2019b), in which the model speaker was provided with prototypical examples of the three Salerno Italian question tunes and was asked to reproduce them in combination with a set of 24 items, all of which were uttered with positive polarity. One of these items (*Stai giocando a domino?* 'Are you playing dominoes?') was used to create the stimuli for the experiment reported here. Specifically we selected two different tonal renditions of this item as base tunes for the stimuli manipulations, i.e., a rise-fall-rise, noted as L+H* HL-H%; and a rise-fall, i.e., L+H* HL-L% (hereafter we will refer to these two tunes as Base tunes). Note that both of the Base tunes shared the same nuclear pitch accent (L+H*), which was located on the last word of the utterance (*domino*) with no prenuclear accents, so that the only tonal difference between the two bases was the presence of a different boundary tone (H% or L%). Resynthesized stimuli were created using PSOLA in Praat (Boersma & Weenink, 2016). Pitch span manipulations were made both within the pitch accent region and the boundary tone. Specifically, three pitch span steps were created by manipulating the height of the peak within the pitch accent and six boundary steps were created to manipulate boundary tone span. The same types of manipulation were made on each base stimulus, after performing a pitch stylization. **Figure 4** shows a schematization of the set of stimuli presented to the listeners.

Reference levels for pitch height and alignment of the targets were taken from the values of a set of 48 utterances produced by the model speaker. Specifically, we set f0 values for 1) the beginning of the utterance (F0i), 2) the L leading tone for the L+H* pitch accent (L1), 3) the pitch accent H peak (PA), 4) the beginning of the LH- phrase accent following the accentual rise (L2), and 5) the height of the boundary tone (BT), while interpolation between these targets was linear. Temporal alignment values for L1, H, and L2 were also based on mean values for the 48 utterances and were held constant across the stimuli. F0i, L1, and L2 were also held constant in height, while H peak height in the pitch accent
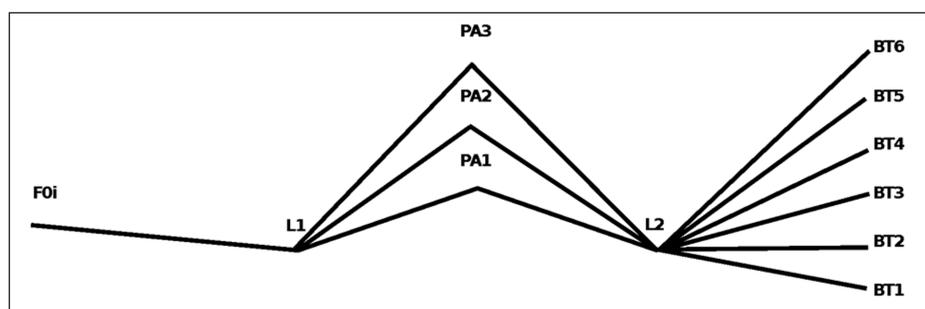

Figure 4: Schematization of the manipulations for the audio stimuli used for the experiment.

(PA) and the Boundary Tone (BT) were manipulated. **Table 1** shows f0 values for each of the experimental stimuli.

PA span manipulations were made starting from PA2 level (the intermediate level), whose height was the mean of the values for all L+H* produced in the 48 reference stimuli. PA1 and PA3 were created by, respectively, lowering and raising the height of the peak by a 15 Hz step. Manipulations for BT were made by using mean values for all L% f0 levels (corresponding to BT2) and by lowering (BT1) or raising (BT3 – BT6) the height of the boundary by 15 Hz. The final number of experimental stimuli was 36 (three PA × six BT × two Base tunes).

4.2. Participants and procedure

Forty-five SI listeners were recruited among the students of the University of Salerno to take part in the experiment. They were roughly balanced across gender (21 males) and age (from 18 to 29 years old, mean: 22.45, median: 21). All of them were born in the Salerno area, where they had also lived for most of their lives. All of them were also currently living in Salerno at the time of the experiment.

The experiment was performed at the linguistic laboratory of the Department of Humanities at the University of Salerno. The test was created and administered to the participants using the research platform Open Sesame (Mathôt, Schreij, & Theeuwes, 2012), which participants accessed using a desktop computer. The experiment started with an instruction slide, after which a test trial would start. Participants had been warmly encouraged to ask questions during this trial phase in case they had doubts. The first author was always present during the experiments and made sure that the task was clear for every participant. After the test trial, the experimental trials would start. Five randomized experimental blocks were presented to each participant.

The 36 experimental stimuli plus 10 filler stimuli[1] were block randomized by participant. Listeners heard the audio stimuli via head-mounted headphones. For each of the stimuli, listeners had to rate the bias of the speaker towards the expected answer to the question. Specifically, they were asked to rate Certainty relative to the expected response to the question stimuli they listened to, i.e., 'What does the speaker expect as an answer to her

Table 1: Height values for the manipulation steps of the experimental stimuli.

Tone Target	f0 height (Hz)
PA1	164
PA2	179
PA3	194
BT1	128
BT2	143
BT3	156
BT4	170
BT5	185
BT6	199

---

[1] Filler stimuli were created by using a different utterance relative to experimental stimuli (i.e., *Hai spostato i mobili?*, 'Did you move the furniture?'). Additionally, fillers differed from experimental stimuli on the phonological level: They were either uttered with a L*+H HL-H% or a L*+H HL-L%, which can also be used in SI to express a yes-no question (see Orrico et al., 2019a).

question? Yes or No?'. They had to answer this question by clicking with the left button of their mouse on any point on a slider going from zero to 100, where zero corresponded to 'The speaker expects NO,' 100 corresponded to 'The speaker expects YES,' and 50 corresponded to 'The speaker does not have any expectation about the answer.' We will refer to this rating as *Certainty score.*

After the test, participants were administered a questionnaire to be filled with personal information. They were asked their names, gender, age, city/country where their parents (or care-takers) were from, and detailed information about in which cities or countries they had lived. In particular we were interested in gathering information about whether or not they had been exposed to languages or dialects other than Salerno Italian. According to the results to the questionnaire, participants were subdivided into two groups, i.e., those who had only been exposed to Salerno Italian and those who experienced exposure to other phonological systems (either other Italian varieties or other languages).

The exposed group included participants who either lived in other cities for prolonged periods of time (longer than 12 months) or who were brought up by non-Salerno Italian speakers. Out of the 45 listeners, 21 had experienced exposure to other phonological systems. Eight participants experienced exposure only by living in other cities, 10 had at least one parent who was not native of Salerno Italian, and three underwent both types of exposure. Among the 11 total participants who lived in other cities, eight had lived in other Southern Italian cities, one in Rome, one in Spain, and one in France. All of the listeners who had lived in other cities experienced such exposure during adulthood. Finally, out of the 13 participants who were raised by speakers of other varieties, 10 had at least one parent from other cities in Southern Italy, one from Venezuela, one from Switzerland, and one from the Republic of Mauritius.

After completing the sociolinguistic questionnaire, participants were asked to complete the Empathy Quotient (EQ) test by Baron-Cohen and Wheelwright (2004), which is a self-report measure of an individual's degree of empathy (both emotional and cognitive empathy). Specifically, it is a 60-item test (40 experimental and 20 fillers), which the participant has to react to by giving a response using a four-point scale ranging from 'strongly agree' to 'strongly disagree.' Specifically, the test allows collecting information about an individual's empathy level, which is the individual's appreciation of others' affective and epistemic states (i.e., perspective taking). For each item within the test, participants can receive one or two points for empathic answers (depending on the strength of the response) or zero for non-empathic answers. Items are roughly balanced according to whether agreeing or disagreeing to the statement implies an empathic answer. Participants, hence, can obtain a score ranging from zero (lowest) to 80 (highest). EQ scores for our participants ranged from 24 to 71 (median: 46).

4.3. Statistical analysis

Data were analyzed using linear mixed-effect regression models fitted in R (R Core Team, 2019) by using the mixed function contained within the package afex (Singmann, Bolker, Westfall, Aust, & Ben-Shachar, 2019) and the Satterthwaite approximation for *p*-values computation. Certainty scores were set as the dependent variable (numerical). Pitch Accent span (PA, 3 levels) Boundary Tone span (BT, 6 levels), Base Tune (Base, 2 levels), and Block (5 levels) were set as main effects. The comprehensive model included also Exposure (listeners' exposure to other phonological systems, 2 levels, i.e., either exposed or not) and participants' Empathy Quotient (EQ, 2 levels). In particular, levels within the variable Exposure were set to 'Yes' or 'No,' depending on whether listeners had long-term exposure to other systems or not. As for EQ, listeners were divided into two groups, i.e., 'High EQ' and 'Low EQ' according to the score obtained in the EQ test. The cut-off

value for the two EQ groups was the median of the scores obtained by all the listeners (46 points). Interactions among all the fixed effects were also included.

The random effect structure was set by following the instructions by Barr, Levy, Scheepers, and Tily (2013), suggesting that a model should contain the maximal random effect structure justified by the design. Specifically, the maximal structure for our design included random intercepts by Listener and random slopes for PA, BT, Base, and Block. Given that the maximal random effect structure did not allow the model to converge, we simplified the structure by eliminating slopes for Base and Block. Inspections of the levels within significant main effects and interactions were made using multiple pairwise comparisons using the functions emmeans::emmeans (Lenth, 2019) and multcomp::as.glht (Hothorn, Bretz, & Westfall, 2008). *P*-values for pairwise contrasts were computed using the Bonferroni method.

Finally, we calculated the coefficients of response variation, which is a standardized measure of the dispersion of a frequency distribution, using the cv function under the R package goeveg (Goral & Schellenberg, 2018) and tested the equality of these coefficients across groups of listeners using the Asymptotic test within the package cvequality (Marwick & Krishnamoorthy, 2019).

4.4. Predictions

In line with the hypotheses outlined in Section 3 above, we formulated the following predictions. Firstly, we predicted that higher steps of both the PA and BT continuum would be found to be significant predictors of the dependent variable (Certainty score), and that higher steps would induce higher positive bias identification. Specifically, for both independent variables, we expected that a progressively higher manipulation step would be interpreted by listeners as being linked to a significantly higher degree of Certainty score.

We additionally predicted that both individual variables, Exposure and EQ, would interact with the main effects of PA and BT span modification. In particular, since we hypothesized that EQ would modulate the relative ability to map pitch information to pragmatic cues, we expected that high-empathy listeners would be more sensitive to both PA and BT manipulations and, furthermore, this would result in a more gradient effect on the Certainty score. As for the role of Exposure, we predicted a wider dispersion of the responses given by Exposed listeners, which would reflect the potential inconsistencies received in the input for specific intonation-meaning mappings as well as a potentially wider exemplar space for mapping the acoustic signal.

Finally, we predicted that the two variables at the listener level (EQ and Exposure) would interact with each other in significant way. We specifically expected that modifications brought about by Exposure (both in terms of response dispersion and in terms of differences in span-meaning mapping) would be higher for high-empathy listeners, as expected from previous literature (e.g., Estève-Gibert et al. 2020).

5. Results

Figure 5 below reports a plot of the Certainty score ratings as an effect of both Pitch Accent and Boundary Tone manipulation. In general, our results are in line with our predictions regarding PA and BT steps. As we can see in **Figure 5**, the main effect of PA shows that wider span induced higher Certainty values, as predicted. Moreover, BT span increase also induced globally greater values of Certainty. The statistical model showed in fact that both PA (F(2, 5788.43): 88.98, $p < .0001$) and BT (F(5, 59.13): 8.38, $p < .0001$) are significant predictors of perceived bias ratings. Crucially, pairwise comparisons among PA levels resulted in all three steps being significantly different from each other, pointing towards a gradual interpretation of the PA height in terms of
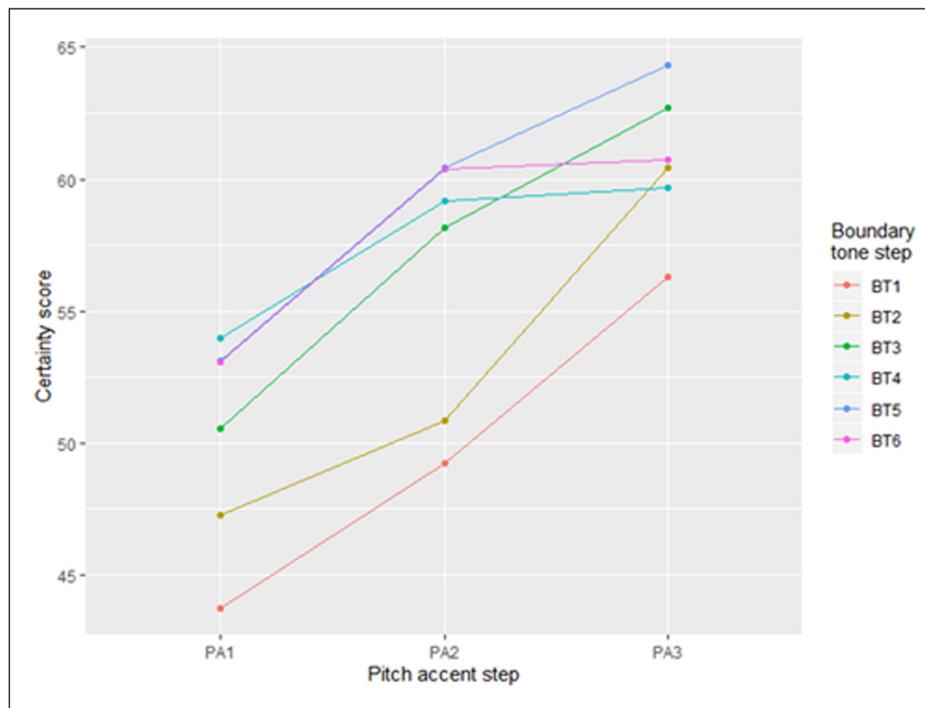
Figure 5: Line plot of the interaction between PA and BT steps in the assignment of Certainty score.

gradually positive bias expectation. Different from pitch accents, pairwise comparisons within BT levels showed that only the two lowest steps (BT1 and BT2) were significantly different relative to the higher ones (BT3 – BT6). Additionally, the interaction between PA and BT reached significance (F(10, 5788.54): 1.87, p = .04). Clearly, the interaction did not affect the direction of the effect, rather its size. Specifically, the combination of the lowest PA (PA1) step with the lowest boundary steps (BT1 and BT2) induced the lowest Certainty ratings (below 50), while higher PA and BT combinations were rated with higher levels of Certainty. Hence, PA and BT span manipulation resulted in an additive effect.

A further effect that reached significance was stimulus Base (F(1, 5790.18): 33.60, p < .0001). Specifically, all stimuli that were created using the H% base stimulus were globally assigned higher Certainty scores (Estimate: 3.74, *SE*: 0.65, z: 5.796, p < .0001) than those created using the L% Base, as visible in **Figure 6**.

Additionally, the effect of Base also reached significance in interaction with PA (F(2, 5792.10): 8.91, p < .0001). Specifically, while for PA3 the difference between the two base tunes was hardly detectable and did not reach significance (Estimate: 1.33; *SE*: 1.12; z: 1.19; p = 0.23), when in combination with PA2 and PA1 such differences were greater and statistically significant (PA1: Estimate: 7.53; *SE*: 1.11; z: 6.78; p < .00001; PA2: Estimate: 2.35; *SE*: 1.13; z: 2.09; p = .04). Note also that the difference between PA levels was stronger for L% Base stimuli, which was confirmed by the estimate coefficients reported in the **Table 2** below.

5.1. Exposure and Empathy Quotient interaction
Against our predictions for H2, our retained model did not show significant main effects of either Exposure or EQ (Exposure: F(1, 41.51): 1.21, p = .28 and EQ: F(1, 41.51): 2.59, p = .12). Nevertheless, and crucially, both individual factors interacted significantly with either PA or BT span step, showing that they do have a role in modulating the mapping of phonetic cues onto perceived bias.
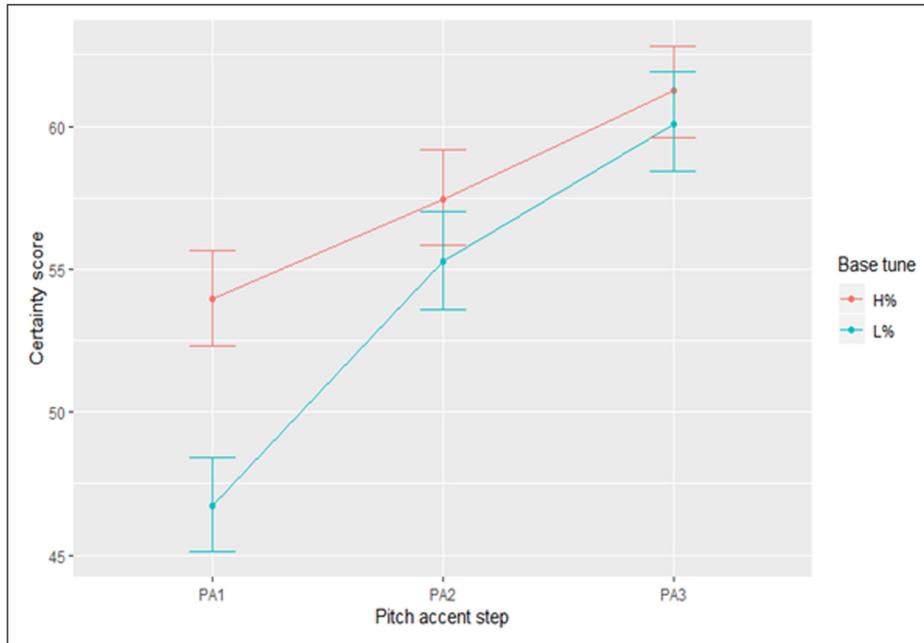
Figure 6: Line plot of the interaction between PA and Base tune in the assignment of Certainty scores.

Table 2: Summary of the pairwise comparisons of PA levels in the two different conditions of Base tune.

	Base: H%				Base: L%			
	Estimate	**SE**	**z. value**	**p. value**	**Estimate**	**SE**	**z. value**	**p. value**
PA1–PA2	−4.05	1.19	−3.38	.002	−9.23	1.19	−7.78	.0001
PA1–PA3	−7.80	1.20	−6.50	.001	−13.57	1.18	−11.52	.0001
PA2–PA3	−3.75	1.19	−3.12	.005	−4.34	1.19	−3.66	.0008

A first clear effect was registered in how EQ scores interacted with BT steps. Note from **Figure 7** that while the High EQ group differentiated across BT steps (red line in the plot), the Low EQ group failed to do so. The interaction between EQ score and Boundary tone was in fact significant (F(5, 59.13): 3.54, p = .007). Specifically, recall that **Figure 5** above showed that boundary levels were a significant predictor of Certainty score ratings, with higher levels of BT being rated as conveying higher degrees of perceived Certainty (for a positive answer). Interestingly, this effect appears to be driven by high-empathy listeners, as revealed by the data in **Figure 7** above. In fact, low-empathy listeners did not make any differentiation in the perception of bias across BT levels, given that none of the paired contrasts reached significance. The high-empathy group, on the other hand, clearly differentiated across boundary types, with significant contrasts between the lowest levels (BT1 and BT2) and the highest ones. The only exception is represented by the highest boundary level (BT6), which did not reach significance in interaction with any of the other levels. **Table 3** shows the details of the multiple comparisons across boundary tone levels for the high EQ listeners group.

The interaction between Exposure and PA levels was also significant, as predicted by H2 (F(2, 5788.43): 6.58, p < .001). **Figure 8** reports a plot with the interaction. Post-hoc analyses showed that listeners who were exposed to non-native phonological systems could not reliably classify the three PA levels as carrying systematically different degrees of perceived bias. Specifically, the analysis of their ratings showed that they only reliably
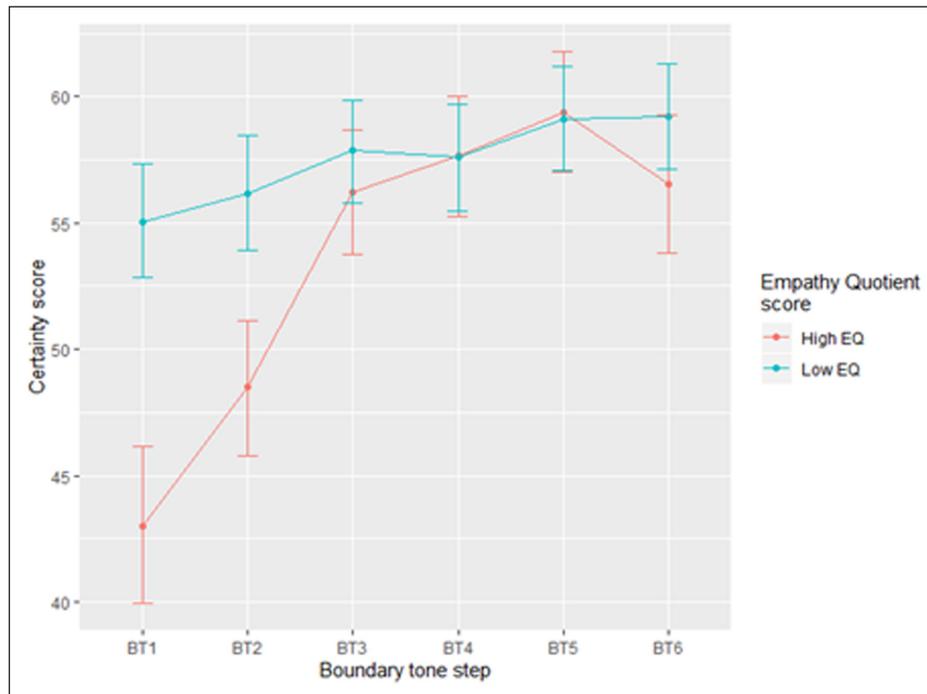
Figure 7: Line plot of the interaction between BT height steps and EQ score in the assignment of Certainty values.

Table 3: Summary of the pairwise comparisons of BT levels for the High-EQ group (left) and the Low-EQ group (right).

	High-EQ				Low-EQ			
	Estimate	**SE**	**z. value**	**p. value**	**Estimate**	**SE**	**z. value**	**p. value**
BT1–BT2	−5.89	3.44	−1.71	.48	−1.04	3.06	−0.34	1.00
BT1–BT3	−13.94	3.23	−4.32	.001	−3.19	2.88	−1.11	.86
BT1–BT4	−14.98	3.46	−4.33	.001	−2.77	3.08	−0.90	.94
BT1–BT5	−17.37	3.35	−5.19	.001	−4.14	2.99	−1.39	.70
BT1–BT6	−14.23	6.06	−2.35	.15	−5.02	5.41	−0.93	.93
BT2–BT3	−8.05	2.20	−3.66	.003	−2.16	1.94	−1.11	.86
BT2–BT4	−9.09	2.52	−3.60	.004	−1.74	2.23	−0.78	.96
BT2–BT5	−11.48	2.37	−4.85	.001	−3.10	2.10	−1.48	.64
BT2–BT6	−8.34	4.46	−1.87	.38	−3.98	3.98	−1.00	.90
BT3–BT4	−1.04	2.22	−0.47	1.00	0.42	1.98	0.21	1.00
BT3–BT5	−3.44	2.04	−1.68	.50	−0.95	1.83	−0.52	.99
BT3–BT6	−0.29	3.98	−0.07	1.00	−1.83	3.55	−0.51	.99
BT4–BT5	−2.39	2.39	−1.00	.90	−1.37	2.13	−0.64	.99
BT4–BT6	0.75	4.51	0.17	1.00	−2.25	4.02	−0.56	.99
BT5–BT6	3.15	4.25	0.74	.97	−0.88	3.79	−0.23	1.00

discriminated PA1 from the other two levels (PA2: Estimate: −7.89; *SE*: 1.16; z: −6.76; $p < .00001$; PA3: Estimate: −9.34; *SE*: 1.16; z: −7.99; $p < .00001$), while assigning non-significantly different degrees of Certainty to PA2 and PA3 (Estimate: −1.44; *SE*: 1.17; z: −1.22; $p < .43$). On the other hand, listeners who were only exposed to their native
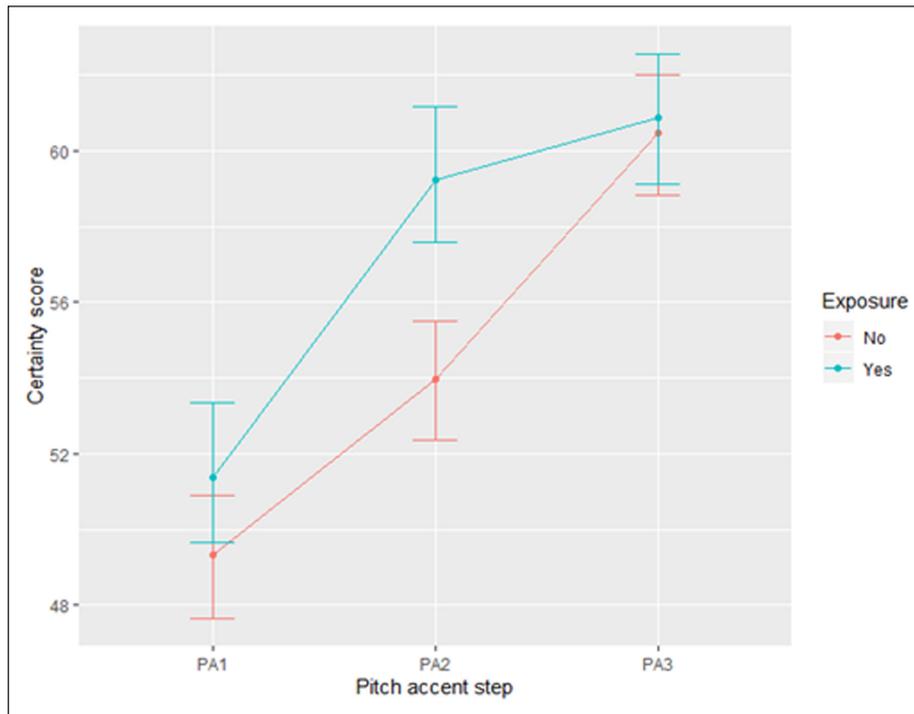
Figure 8: Line plot of the interaction between PA height steps and Exposure in the assignment of Certainty values.

Table 4: Summary of the pairwise comparisons of PA levels for the High-EQ (H-EQ) and Low-EQ (L-EQ) groups in the two Exposure conditions.

		Exposure: No				Exposure: Yes			
		Estimate	SE	z. value	p. value	Estimate	SE	z. value	p. value
H-EQ	PA1–PA2	−5.458	1.597	−3.417	.002	−9.042	1.752	−5.160	<.001
	PA1–PA3	−15.547	1.585	−9.810	.001	−8.398	1.743	−4.819	<.001
	PA2–PA3	−10.089	1.599	−6.310	.001	.644	1.747	.368	.928
L-EQ	PA1–PA2	−3.394	1.415	−2.399	.043	−6.753	1.544	−4.373	.001
	PA1–PA3	−7.601	1.404	−5.415	.001	−10.277	1.555	−6.610	.001
	PA2–PA3	−4.206	1.406	−2.991	.008	−3.523	1.566	−2.250	.063

variety showed a consistent rating of the three different pitch accent levels as conveying different degrees of Certainty, with significant results for all pairwise comparisons (PA1-PA2: Estimate: −4.43; *SE*: 1.06; z: −4.14; $p < .001$; PA1-PA3: Estimate: −11.57; *SE*: 1.05; z: −10.93; $p < .00001$; Estimate: −7.14; *SE*: 1.06; z: −6.71; $p < .00001$).

In line with the predictions, when considering the three-way interaction with EQ, the differences between the two Exposure groups appear to be even more prominent. The three-way interaction among PA, Exposure, and EQ in fact reached significance (F(2, 5788.43): 6.60, $p < .001$). The plot in **Figure 9** shows the relationship among these three variables, while **Table 4** reports the details of the pairwise comparisons.

Clearly, listeners with exposure to non-native systems failed to associate different degrees of Certainty to the PA2-PA3 pair, regardless of their EQ levels. However, as shown in **Figure 9**, high-empathy listeners with no exposure to non-native dialects made a clear differentiation across the three PA levels, with the narrowest pitch span being associated with a negative bias and the widest pitch span to a positive bias. Low-empathy listeners within the same Exposure group, on the other hand, rated Certainty for all
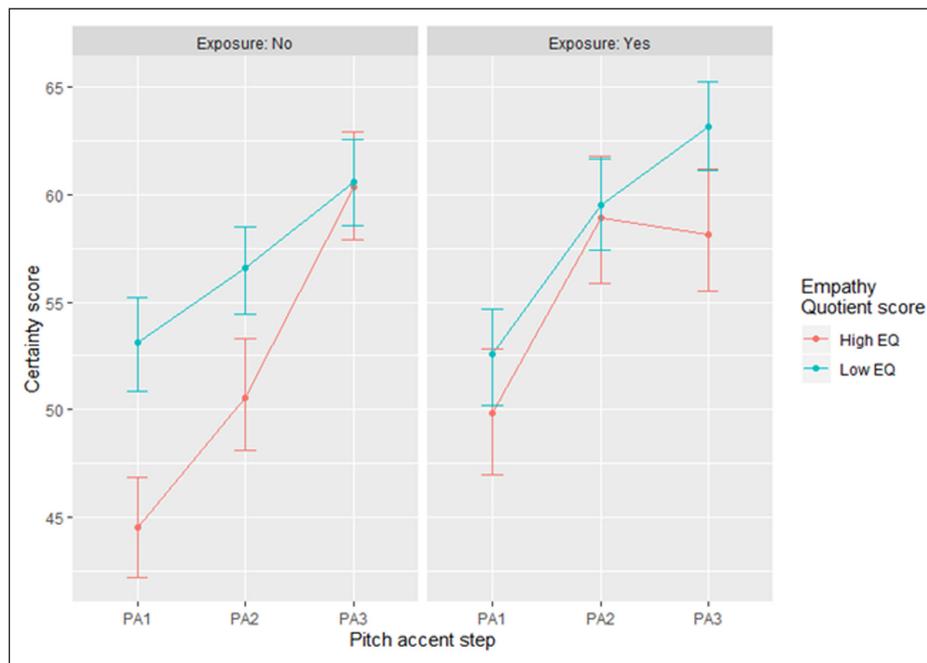
Figure 9: Line plot of the interaction between PA height steps, Exposure, and EQ score in the assignment of Certainty values.
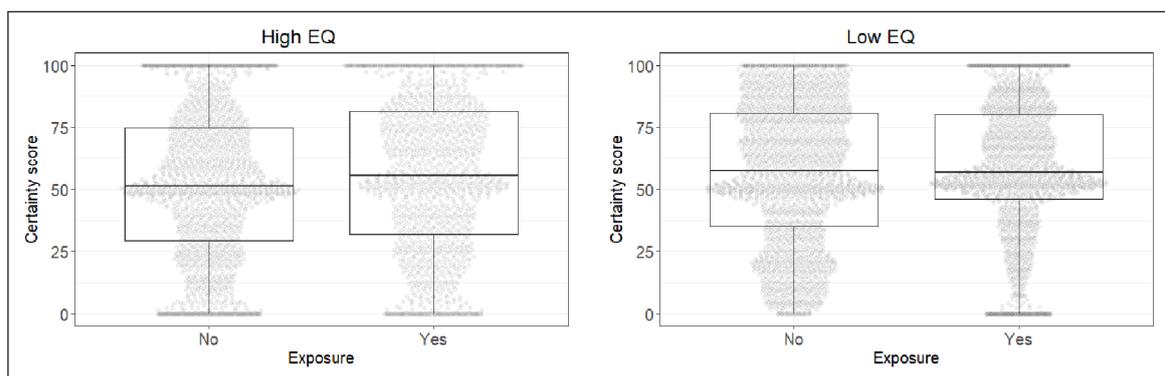

Figure 10: Boxplots of the distribution of the assignment of Certainty values for High EQ (left) and Low EQ (right) listeners in the two Exposure groups (x axis).

the PA levels, on average, above 50, though maintaining a significant contrast among them. Additionally, the plot also shows a tendency for low-empathy listeners to behave in a very similar fashion in the two Exposure conditions, while listeners with higher EQ scores showed a stronger contrast between the Exposure conditions. The interaction between Exposure and PA level only reached significance for PA2 (Estimate: –8.21; *SE*: 4.05; *z*: –2.02, $p = .043$).

Finally, in order to address the specific hypothesis that Exposure would affect the variability of responses, we compared coefficients of variation within Certainty score ratings given by the two Exposure groups. Results of the Asymptotic test yielded significant results (test statistics: 5.327, $p = .021$), though, contrary to predictions, non-exposed listeners yielded higher coefficients of variance (0.54) with respect to exposed listeners (0.51). However, visual inspection of the dispersion of response data suggests an interaction between Exposure and EQ. The boxplots in **Figure 10** show that while a higher variation of responses was found only for the non-exposed High EQ listeners, a smaller variation was registered for low-empathy exposed listeners (**Figure 10**, right panel).

Comparisons of equality of variation distributions between the two EQ groups of the exposed listeners yielded indeed significant results (test statistics: 30.203; p < .0001), showing that EQ score modulated the variance of responses within the exposed group.

6. Discussion
6.1. Summary of results
In a perception study, we investigated the role of pitch span within the nuclear configuration region of Salerno Italian yes-no questions in the recovery of a speaker's epistemic bias. Two different base tunes (L+H* HL-L% and L+H* HL-H%), differing in the direction of the boundary tone (rising or falling) were stylized then resynthesized with manipulated pitch span (gradually larger f0 excursion between L and H targets) for both pitch accent and boundary tone region and were then presented to listeners, who evaluated each stimulus in terms of perceived speaker Certainty about a 'yes' response. Results show that pitch span has a crucial role in listeners' identification of speakers' epistemic bias and, in the specific case tested here, in evaluating whether the speaker expects a positive answer to a question.

The general effect that was detected is in line with our main hypothesis (H1), which is that a wider span in both rising pitch accent and rising boundary tone region would convey a higher degree of a speaker's positive bias (hypothesis borne out of observation of SI production patterns), hence resulting in higher Certainty scores. What is more, the separate contributions of pitch accent span on one side and of boundary tone span on the other appear to have an additive effect, so that the highest Certainty scores were recorded for combinations of higher PA and BT steps. Additionally, the effect of variability at the listeners' level was investigated and, crucially, results point towards a combined effect of both the cognitive (EQ) and our sociolinguistic (Exposure) variables, as predicted by H2. Specifically, the significance of these variables was found both in interaction with tonal cues and with each other, showing that they significantly and gradually affect the way epistemic bias is extracted from intonational cues. Additionally, we found a difference in the variation of responses between the two Exposure groups, which was crucially modulated by the EQ score. While this specific result does not directly confirm our predictions (higher response variation for exposed listeners), it provides evidence for the interaction between Exposure and EQ. In general, these results strongly suggest that epistemic bias identification is not homogeneous throughout the listeners' community, and that individual cognitive skills interact with regional dialect exposure effects.

6.2. Gradience of cues and meaning
First, the findings reported here add to the literature on biased questions and intonation. A link had already been suggested also for other languages (e.g., Prieto & Borràs-Comes, 2018 for Catalan) and for other varieties of Italian (Savino & Grice, 2011 for Bari Italian) given that, as reported in the introduction, intonation is argued to have a prominent role in the production and perception of epistemic bias in yes-no questions. Our results contribute to this literature by uncovering the crucial contribution of gradual pitch span variation to perceived epistemic bias. As for the pitch accent region, Certainty scores were gradually affected by the step manipulation, suggesting that gradient pitch cues do not necessarily need to map onto discrete meaning categories. Hence, an important contribution of these findings is that gradual intonation cues map onto gradient meanings of pragmatic nature, beyond traditional phonetics/phonology mapping. As for the boundary manipulation, our results suggest a stronger mediation of the phonological level, in that only the lowest and narrower stimuli (BT1 and BT2) behaved differently from all the others.

As reported in the introduction, most studies of intonational meaning build on the assumption that specific intonational categories align with specific expression of speakers' epistemic dispositions in a dichotomous way, e.g., Bartels' (1999) L-/H- opposition mapping onto the presence/absence of an 'assert morpheme' or Gunlogson's (2004) rise/fall opposition mapping onto presence/absence of speaker commitment. Nevertheless, recent contributions within Conversational Analysis have proposed that epistemic disposition can be better modeled in a gradient rather than a categorical way. Heritage (2012), for example, has proposed that English can encode several steps along a continuum reflecting gradient epistemic stance. For example, interrogatives (Are you married?), tag-questions (You're married, aren't you?), and declaratives (You are married.) would encode three different steps from an 'unknowing' to a 'knowing' epistemic stance relative to a specific matter, i.e., the proposition 'You are married.' Additionally, several linguistic strategies might be used to encode intermediate degrees of epistemic stance and, crucially, these strategies might not be interchangeable. Oshima (2017), for example, proposes that common strategies used in English to express a biased question (e.g., negative polarity and tags) carry information regarding the relative degree of bias. Specifically, tag questions might convey a higher degree of positive epistemic bias relative to negative polar questions. Our perception results can be interpreted in light of these proposals, though by assigning a stronger contribution to intonational cues, which alone would map onto several steps along a continuum of epistemic stance.

This provides further evidence against the traditional view of pitch range variation as only conveying paralinguistic meanings (as in Liberman & Pierrehumbert's [1984] 'free range hypothesis'). To the best of our knowledge, no intonational meaning model has attempted to formalize the contribution of gradual phonetic cues in the expression of meaning. The only model that assigns specific importance to relative pitch height/span information is Gussenhoven's (2002; 2016) Biological Code theory, in which it is predicted that the rules modulating the meanings of intonational categories are yielded by grammaticalizations of universal meanings assigned to f0 (and intensity) modulations. Crucially, however, the gradient side of the form-meaning mapping in Gussenhoven's theory is limited to the expression of paralinguistic meaning. For example, according to the Frequency Code, the paradigmatic opposition between high and low fundamental frequency (or rising and falling movements) would universally signal, respectively, an opposition between submissiveness and/or uncertainty on one side and confidence and/or finality on the other. On the grammatical side, this paralinguistic opposition would give rise to the opposition between questions and statements, modulated by language-specific constraints (see Chen et al., 2004). Similarly, according to this view, an expanded rise span would only signal non-linguistic meanings, such as polite attitude. The Effort Code, in a similar fashion, would predict that wider pitch movements would universally signal emphasis or surprise.

The findings reported here, however, do not easily fit into such a universal way of deriving the meaning of intonation. First of all, the gradience that we report is pragmatic in nature and therefore differs from Gussenhoven's predictions that paralinguistic gradience is grammaticalized in distinct linguistic categories. Additionally, the oppositions found between the way high versus low pitch cues are mapped onto bias can hardly be explained as deriving from such universals. The clearest example is given by the Certainty ratings resulting from boundary tone steps. On the basis of predictions derived from the Frequency Code, we would have expected that higher BT steps would have yielded lower Certainty scores, while what we found was the opposite. As mentioned in the introduction, several predictions of the Frequency Code have not found support in recent experimental research

investigating the role of pitch height in the expression of deference and politeness in Korean (Brown et al., 2014) and in Russian (Chikulaeva & D'Imperio, 2018). As for Korean, it appears that politeness is mainly conveyed by loudness variation and not pitch (Idemaru, Winter, & Brown, 2020), rendering the idea of universality of paralinguistic use of gradual pitch cues rather problematic. Also, incredulity meanings have often been linked to expanded pitch span (e.g., Hirschberg & Ward, 1992; Savino & Grice, 2011; Borràs-Comes et al., 2014), which could easily be interpreted as a grammaticalization of the Effort code. This is not in line with our findings, though, in that the link found here between narrower span in pitch accents and negative bias appears to go against the universality of intonational meaning predicted by this code. Therefore, our results point to the need to formulate language- and dialect-specific mappings between gradual acoustic-intonation cues and gradual pragmatic meanings.

The gradient relationship between gradient cues and meaning resulting from our study would be compatible with recent experimental evidence reported by Holliday and Villareal (2020), who test the social meaning of intonation. Specifically, the study investigates the link between both phonological elements (pitch accents) and phonetic detail (continuous pitch span and voice quality information) and social meaning related to speaker blackness. The findings of this study point towards meaningful variability of both phonological and phonetic information, with an incremental effect of continuous phonetic cues. Furthermore, the authors also find that the meaning of gradient cues is not shared across phonological categories. While for L+H* (phonologically considered more frequent in African American English than H*) gradient span contributed to assess a stronger social meaning related to blackness, the same was not found for H* (though results for phonological manipulation of the pitch accent did not show different listener behaviors). This suggests that phonological and phonetic properties of intonation work together in determining aspects of intonational meaning and, crucially, that variable meanings retrievable from intonational allomorphs might arise from modifications of the core meaning of that category.

One crucial difference, however, between the treatment proposed by Holliday and Villareal (2020) and our own is in the type of gradient meaning that maps onto the continuous cues. While Holliday and Villareal consider social meaning linked to the ethnicity of the speaker, we focused on dialectal variability. Nevertheless, a similar relationship between phonology and phonetics in the construction of meaning might be applicable to our data as well. While we did not specifically test for the relationship between phonological and phonetic elements (e.g., meaning of span modifications in two phonologically different pitch accents), we can speculate on the mediation effect provided by the phonological category.

Independent evidence for SI about the meaning of span modifications in different pitch accent categories has already been reported in Orrico and D'Imperio (2020) as far as wh-questions are concerned. Specifically, the study reports that a greater span in L*+H accents is linked to counter-expectational meanings (in terms of rejection of a previously uttered proposition), which has been analyzed as a lack of speaker commitment to a relevant proposition. On the other hand, the results reported in the present study appear to assign the opposite meaning to a greater span, which would suggest some degree of mediation of the phonological form. Similar results were reported by Chikulaeva and D'Imperio (2018) for the tonal encoding of politeness in Russian. Specifically, they report data from a production study in which f0 height in pitch accents positively correlated with the expression of politeness. Crucially, however, the same effect was not found for all pitch accents categories (e.g., different from H*, H+!H* accents were not found to modulate politeness as a function of f0 height), providing additional evidence

of the interaction between phonology and phonetics of intonation in the expression of meaning. We, hence, suggest that while intonational phonological units might possess meanings, modifications at the sub-phonemic (more precisely sub-morphemic) level might be employed to modulate those core meanings. This treatment of intonational meaning and, above all, of the way phonological and phonetic elements cooperate to the construction of the global meaning is reminiscent of the proposal made by Gili Fivela (2008) for Pisa Italian (a central variety), in which it is argued that different 'shades of meaning' of a given phonological category can be conveyed by phonetic modifications within that category.

The meaningful character of phonetic modifications of a phonological category, however, does not appear to apply to all the structural cues of an intonational contour. As a matter of fact, such behavior of the phonology-phonetic interplay in the transmission of meaning emerges, in our data, only to pitch accents, while the same cannot be said for boundary tones. Boundary tone level was in fact found to bear intrinsic meanings related to the expression of question epistemic bias, since finer span modifications within the two broad phonological categories (L% and H%) did not appear to contribute to further characterize meaning. It is possible, though, that continuous modifications of boundary tones can map onto continuous meanings in the same way as pitch accents do, though for other dimensions of meanings which we did not explore here.

Finally, an effect of base tune was also found in our data. Specifically, the stimuli created on the basis of the rise-fall $L+H^*$ HL-L% base stimulus were assigned lower overall Certainty scores. This effect indicates that cues beyond pitch are also exploited to infer epistemic bias in SI. Note that the effect of Base also interacted with PA steps, hence determining how step height was employed for Certainty rating in that region. We do not know how f0 directionality might be related to spectral information or other non-pitch intonational cues, given that this is an area that has received less attention. However, these results are in line with findings suggesting that, for example, local duration cues play a role in the production of sentence modality in Neapolitan Italian (Cangemi & D'Imperio, 2015) adding to the complex relationship between the phonetics-phonology mapping on the one hand, and the phonology-pragmatic meaning mapping on the other.

Taken together, our results provide useful insight on the relationship between intonational cues and pragmatic meanings. First of all, continuous phonetic modifications within a phonological category are not randomly variable, rather they carry specific meanings of pragmatic nature. Such meanings, we posit, are not completely independent from the phonological category they belong to, but they represent modifications of the core meaning of that category. In other words, the phonetic space that an allophonic modification occupies is informative with reference to the strength of the meaning associated with the phonological category. Also, these modifications to the core meaning are incremental, which means that the specific meaning of a given phonological category is modified stepwise along a continuum of gradient modifications of the pitch accent form. In this way, gradient intonational cues at the sub-phonemic level map onto gradient pragmatic meanings.

6.3. Exposure and Empathy interaction

Crucially, our data suggest that the relationship between gradient cues and gradient meanings is not stable within a linguistic community. This is because the way in which each manipulation step is mapped onto a specific degree of bias varies also as a function of listener-specific characteristics at both social (exposure to other systems) and cognitive (degree of empathy) levels. These two types of variable would then operate over two distinct planes for the construction of specific form-meaning mappings.

Exposure—a variable of social nature—operates over the organization of the phonetic space in terms of width and shape of the category space (see also Clopper, 2014 for segmental models). Prolonged exposure to other phonological systems might cause the listener to widen the specific category and, in general, to develop more flexible category boundaries. In other words, listeners who undergo the process of long-term exposure to other dialects or languages are, generally speaking, exposed to much more variable input than those who do not, which is the reason why exposed listeners might attend more closely to fine phonetic information to assess the meaning of a question given the richness and variety in their input. This is reminiscent of findings reported by Chen et al. (2017), in which bi-dialectal listeners (Australian English listeners with at least one parent native of another English dialect) were found to have widened the phonetic space relative to vowel phonemes and, therefore, being much more tolerant to mispronunciations than mono-dialectal ones, though, in our case, the flexible boundaries are built for sub-phonemic categories, rather than phonological ones.

Empathy—a cognitive variable—operates, on the other hand, over the cue-meaning mapping of gradient cue-meaning itself or, in other words, it defines how fine-grained the mapping between intonational cues and meanings is. A listener with a higher degree of empathy would be much more attentive to fine phonetic details within the input and the way this detail is linked to specific (degrees of) meanings. In this way, a high-empathy listener is predicted to be able to access more steps along the cue-meaning continuum. This empathy-induced sensitivity would also explain why high-empathy listeners are much more affected by variable exposure to other systems. Empathy skills would then lead them to attend more closely to phonetic cues of the input, resulting in a faster and stronger perceptual change and causing an augmented sensitivity to category interference. The scenario described above is supported by an analysis reported in **Figure 11**. For the
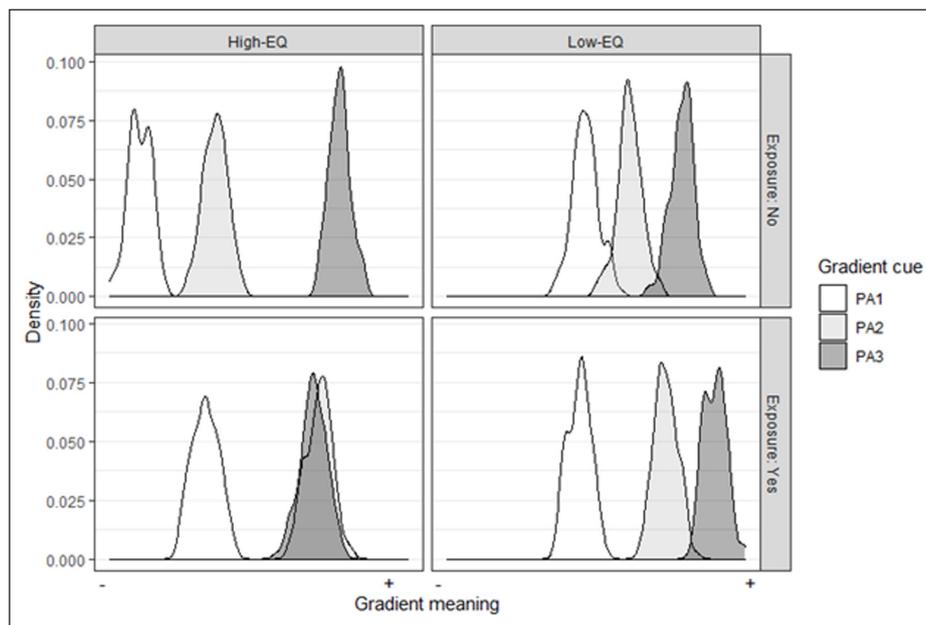

Figure 11: Density plot illustrating the difference in epistemic bias ratings for the three PA pitch accent span steps in the four different groups of listeners: non-exposed/High-EQ (top-left), non-exposed/Low-EQ (top-right), exposed/High-EQ (bottom-left), and exposed/Low-EQ (bottom-right). Different shades refer to different span steps within the L+H* pitch accent, with PA1 indicating the lowest span step and PA3 the highest. While showing that for all listeners gradient span cues can be associated with different degrees of meanings, the plot crucially shows the way in which both Exposure and EQ interact with the way in which intonation-meaning mapping gradient cues-gradient meanings mappings is shaped within the individual grammar.

sake of simplicity, we will restrict this section of the discussion only to our pitch accent manipulation.

The density plots reported in **Figure 11** above show the effect of the interaction of our two variables (Exposure and EQ) on the span manipulation for the PA. The plot was created by using a simulated dataset, based on results obtained in the experiment reported in this paper, reporting therefore the way in which each group rated each pitch accent step. Specifically, the dataset was created by generating random data points, normally distributed around the means of each one of the four participant groups (non-exposed/High-EQ, non-exposed/Low-EQ, exposed/High-EQ, and exposed/Low-EQ). Standard deviation values were reduced relative to the results of the experiment. This was done in order to eliminate some of the noise, coming mainly deriving from the interaction between pitch accent steps and boundary tone steps. However, proportions for standard deviations across pitch accents steps and across the four different groups were the same as the actual dataset resulting from the experiment. As anticipated above, the plot refers to the three pitch accent steps (PA1, PA2, and PA3).

The first thing to be noticed from the plot above is that some degree of gradience is present for all the listeners independent of the specific sub-group they belong to. Recall that the gradient cue modifications for pitch accents did not involve a categorical shift at the phonological level, supporting the idea that it is hence possible to specify meaning of pragmatic nature at a sub-phonemic level. If we focus now on the non-exposed groups (top panels) we see that both high- and low-EQ listeners rated each step independently from each other (given the shape of the distributions), though with a difference in magnitude. Specifically, data for the (non-exposed) high-EQ group reveal the existence of a larger distance in terms of Certainty scores across the three steps. This might be taken as evidence of the existence of meaningful intermediate steps (e.g., between PA2 and PA3), which would point towards a more fine-grained mapping. On the other hand, the exposed groups (bottom panels), and in particular the high-empathy individuals, globally reduced the distance between the two highest steps (PA2 and PA3), indicating that their phonetic space might have been reorganized in relation to the meaning expressed, i.e., by enlarging category space they are specifying the highest degree of meaning as being expressed by a larger sub-phonemic category with reference to non-exposed listeners. In other words, being exposed to a different phonological system would not simply cause them to lose the ability to link gradual modifications of the pitch accent to different degrees of meaning, rather it would affect the way in which some of the manipulation steps are linked to gradual meanings. Similar to the non-exposed group, the role played by EQ is expressed in terms of the magnitude of the effect, with (exposed) high-EQ listeners showing a stronger merge of the two steps, hence pointing again towards a more developed sensitivity of these listeners to elements within the input (which in this case are predicted to be unreliable), leading to a stronger effect of Exposure in terms of native category boundary blurring.

These findings are in line with exemplar approaches to sociolinguistically induced segmental variability, such as those discussed in a study on different levels of exposure to either General or Northern American English (cf. Clopper 2014), proposing that the combined distributions of exposure to both local and standard varieties are expected to be more complex and potentially more overlapping (due to spread of phonemic distributions in the combined dialect language users). As noted by Clopper (2014), various classes of exemplar models can indeed account for simultaneous effects of facilitation, on one side, and interference, on the other, due to multiple levels of linguistic and indexical information which are simultaneously stored and represented. Moreover, exemplar-based models predict that both recency (short-term novel exposure) and frequency (long-term exposure) interfere with lexical processing so that individuals with exposure to lower

levels of sociolinguistic and dialectal variation "should have relatively less variable distributions of phonological and/or lexical exemplars" (Clopper 2014, p. 71). Hence, the dialect exposure effect found in our study appears to mimic what has been already claimed in the segmental/lexical phonology literature, and could be accommodated by similar models designed for intonational phonology purposes.

As reported above, exemplar models build on the idea that detailed information is retrieved from linguistic input and is stored in memory together with contextual information. Exposure, therefore, is deemed to be a key factor in this process, also to navigate the uncertainty of the speech signal (as predicted by inference-based models). The role of individual EQ levels, though, cannot be subsumed by any of the existing models. In other words, extrapolating category distribution from Exposure (both short-term and long-term) cannot be taken to be a uniform process independent of individual cognitive skills. As noticed above, shape and the width of the density plots in **Figure 11** are taken to mirror category distribution though not independently of individual empathy skills. Hence, both language change and immediate processing outcomes, at least as far as the form-meaning mapping is concerned, cannot be entirely predicted by current exemplar models. Though these models allow perception-production interference, they still do not predict category-formation and processing variability that would have a cognitive basis.

The effect of empathy discussed here might be considered as a variable that modulates the sensitivity of the listener to the information that is present in the input and her ability to consolidate the indexing of the variable linguistic form as a function of contextual information. It is not clear at this point, however, at what level this variable operates, whether it affects the mapping between portions of phonetic space with specific meanings (or contextual information), or the partition itself of the phonetic space. Future research should aim at addressing this issue.

6.4. The role of social and cognitive variables in phonological categorization: Which perspectives?

The exact picture of the mapping between prosodic cues and pragmatics is, given our results, rendered even more complex by considering the role of variability at the listener level. Our results show that both listeners' exposure to other phonological systems and their EQ score have a significant impact on the way pitch span is exploited to infer epistemic bias. As for the role of Exposure, in our data this variable appeared to interact mainly with PA step manipulation. In fact, while listeners who were only exposed to their native dialect significantly linked the three different PA steps with different degrees of bias, exposed listeners showed a less gradual and more binary perception behavior (by providing different rating only to the lowest and the highest PA steps). While it is well documented that speech encodes socio-indexical information on the part of the speaker, these results point out the need to consider different kinds of individual variability also at the listener/perceiver level.

At a first look, our findings are easily accounted for by models arguing for the role of past experience (speech input) in phonological representation, as for instance exemplar-based models (Johnson, 1997; Goldinger, 1998; Pierrehumbert, 2001; Todd, Pierrehumbert, & Hay, 2019). The basic idea behind these models is that past linguistic experience is stored in an individual's representation in a very detailed way. This theory has been fueled by experimental findings supporting the idea that mental representation of sound structure is continuously updated and enriched with phonetic details as an effect of exposure to new instances of a given category (Goldinger, 1998; 2000). This view is against the traditional treatment of mental representation of phonology in which continuous, and more crucially variable, phonetic details are treated like meaningless noise and filtered out by listeners

in perception (e.g., Halle, 1985). These two positions are poles apart as far as abstract representation is concerned, with one position seeing the abstract representation as a by-product of the processing of similar sounds, while the other taking the abstract category as the only information that is actually stored.

However, more recent treatments have assumed a hybrid perspective, positing the importance of taking into account an enriched phonological representation (e.g., with community-specific information stored into memory) but, at the same time, by also stressing the importance of abstract representation, needed for instance to account for the perception and production of new words, which cannot be captured in early exemplar models (see Pierrehumbert, 2002; German et al., 2013 for arguments supporting this view). In other words, both abstract phonological information and gradual phonetic detail are stored into memory, and influence both production and perception of upcoming words and utterances.

Similar mechanisms appear to emerge for the intonation-meaning mapping. Recent evidence about this comes from recent work on social implicit priming in intonation reported in Portes and German (2019). As briefly reported in the introduction, the study reports on the way Corsican French listeners process the meaning of a rise-fall tune uttered by Continental French speakers, in terms of question or statement interpretation. The contour tested is a specific rise-fall that is grammatically legal in both varieties, though with different meaning, i.e., question in Corsican and statement in Continental French. Results of this study show that the priming effect, that is evoking either France or Corsica tune-modality mapping, is indeed significant, with listeners exposed to the Corsican prime showing higher question responses. Portes and German (2019) explicitly discuss these results framing them within an exemplar-based model predicting that, as an effect of exposure to the two varieties of French, different form-meaning mappings would be stored in memory, each associated with a specific variety. The role of the regional prime would therefore be that of activating one of the two mappings. Interestingly, the authors posit a great importance of the effect of relative exposure to the two varieties, which would have the role of modulating the strength of the priming effect. More specifically, they argue that a balanced exposure to the two varieties would result in the strongest effect of the prime manipulation. Even though our study does not allow us to make specific predictions about how different mappings between intonational cues and expression of epistemic bias in questions can be activated, our results are in line with the findings of previous studies in the sense that they clearly show that individuals can variably process intonational cues (e.g., the relationship between pitch span and identification of bias) as an effect of having been exposed to non-native phonological systems (a variable that the studies reported above did not directly manipulate).

An alternative model allowing for explaining how phonemic perception can change as an effect of variability in the input has been proposed by Kleinschmidt (2019), in which segmental perception is modeled, using Bayesian rules, as a process of inference under uncertainty. Specifically, the author claims that listeners manage to adapt to speaker-specific variability because each dimension of speaker variation has a given degree of informativity and of utility. These dimensions of variability might be more or less informative with regards to the distribution of cues used to produce a given category, while also being more or less useful for the listener to predict (given particular indexical information of the speaker) the mapping between phonetic cues and phonological categories. In other words, listeners possess prior expectations about linguistic variability, in terms of both what can be variable within the speech signal as well as the source of such variability (e.g., dialect or gender), hence exploiting such knowledge to navigate inconsistencies in

the speech input. We also know that listeners can rapidly adapt to variable intonational input. For instance, Kurumada, Brown, and Tanenhaus (2012) and Tanenhaus, Kurumada, and Brown (2015) report that listeners, in order to navigate variability in the intonational input, are able to rapidly adapt to speaker-specific uses of intonation to encode pragmatic meanings. In other words, they are able to update their expectations about speakers' intended meaning in specific contexts based on past experience. Note though that both inferential and adaptive approaches would need to explain perceiver-specific effects in light of our results.

Apart from the role of socio-indexical priming in determining intonational meaning mapping, also argued for uptalk use in New Zealand English by Warren (2017), the role of cognitive variables might modulate how the perceiver extracts epistemic stance. First of all, it appears that, as also suggested by the literature reviewed in Yu and Zellou (2019), not only socio-indexical information but also listeners' cognitive skills have a role in shaping the way speech is perceived and processed. As for the effect of EQ quotient in our study, it appears that the degree of empathy mainly affects the way listeners interpret tone boundary information, with low EQ listeners failing to employ variable span in boundary tones to the purpose of determining epistemic bias, as opposed to high EQ individuals. In addition, our results also suggest that high EQ individuals might use a wider range of cues to interpret epistemic stance, relative to low EQ ones. This is in line with recent literature arguing for different interpretative strategies in individual listeners, such as the difference between pragmatic responders and semantic responders (Degen & Tanenhaus, 2015), and high-empathy levels being linked to greater (as well as earlier) use of pitch cues in recent Eyetracking data (Estève-Gibert et al., 2020) on intonation processing in French. However, different from the study reported in Estève-Gibert et al. (2020), in which listeners could also rely on later lexical disambiguation, here only intonational cues were available for the interpretation of the speaker's intended meaning. A heightened sensitivity to pitch cues appears to be also true for individuals scoring low in the Autism Spectrum quotient (Jun & Bishop, 2015), which is subsumed by the cognitive component of the EQ.

In our experiment, remember that EQ was found to also have an effect in modulating the degree of the Exposure effect, with higher EQ predicting that the listener would be more affected by exposure to other phonological systems. In other words, while it can be reasonably assumed that exposure leads to the creation of a new form-meaning mapping having a specific contextual label, we should avoid assuming that the same type and length of exposure would have exactly the same effect on all individual listeners. The notion of an ideal listener navigating the many sources of production variability and contextual effects should hence be extended to include perceiver-specific processing styles. This would also apply to inferential Bayesian models accounting for the bidirectional role of socio-indexical and linguistic cues in segmental perception (Kleinschmidt, 2019). In fact, neither current versions of exemplar-based or inference-based models can account for the empathy interaction effects reported here.

Note that in early sociophonetic accounts of indexical variability, a dichotomy was introduced to distinguish between speaker versus 'talker' normalization (Johnson, 1990; Johnson et al., 1999), to underline the fact that phonemic categorization is guided by individual talker variability that goes beyond mere speaker (i.e., anatomically based) differentiation. Analogously, our data point to the role of an individual 'perceiver,' endowed with specific cognitive skills, shaping category retrieval of intonational form and meaning. In other words, indexically based talker variability is not sufficient to account for exemplar space structure in guiding phonological processing. Hence, perceiver-specific processes need to be integrated with more general listener processes in accounting for

phonological processing. The notion of an ideal listener navigating the many sources of production variability and contextual effects should be accompanied by perceiver-specific processing styles. This applies also to inferential Bayesian models trying to account for the bidirectional role of socio-indexical and linguistic cues in segmental perception (Kleinschmidt, 2019).

Hence, independent of the specific perception model that is adopted, be it exemplar or inferential, our claim is that we cannot simply assume that listeners' strategies in making sense of variability in the speech signal would be necessarily homogeneous across a linguistic community. Future work should take into account different sources of individual variability and its impact on the intonation-meaning mapping as well as on phonemic activation, possibly by postulating that cognitive, perceiver-based specificities would both modulate socio-indexical priming effects as well as allow or inhibit the impact of different kinds of dialectal exposure.

7. Conclusion

We here report the results of a perception experiment on the effect of pitch span manipulations on perceived degree of epistemic bias in Salerno Italian yes-no questions. Our data show a gradual mapping between phonetic cues on one side and epistemic stance measured through Certainty scores, though details of this mapping appear to be sensitive to both long-term exposure to non-native dialects as well as to a cognitive variable (Empathy). Specifically, we found a general effect of wider pitch span in signaling greater positive bias, in both pitch accent and boundary tone step-manipulations, with an additive effect. Additionally, the acoustic-meaning mapping was heavily affected by cognitive individual variability, here defined by our use of Empathy Quotient (EQ) scores, and in particular in the way long-term exposure to non-native dialects and other languages affected bias identification. Specifically, EQ interacted with exposure levels in determining both the details of the sound-meaning mapping as well as the magnitude of the pitch span effect. The results support a view in which both social and cognitive variables affect the path from acoustic-intonational cues to gradual pragmatic meaning processing and hence call for perception models taking into account individual perceiver behavior.

Acknowledgements

Competing Interests

The authors have no competing interests to declare.

References

Armstrong, M. E. (2012). *The development of yes-no question intonation in Puerto Rican Spanish*. Doctoral dissertation, The Ohio State University.

Arvaniti, A., Baltazani, M., & **Gryllia, S.** (2014). The pragmatic interpretation of intonation in Greek wh-questions. *Speech Prosody 2014* (pp. 1144–1148). DOI: https://doi.org/10.21437/SpeechProsody.2014-218

Asher, N., & **Reese, B.** (2005). Negative bias in polar questions. In E. Maier, C. Bary & J. Huitink (Eds.), *Proceedings of Sinn und Bedeutung 9 (SuB9)*. www.ru.nl/ncs/sub9

Asher, N., & **Reese, B.** (2007). Intonation and discourse: Biased questions. *Interdisciplinary studies on information structure, 8,* 1–38.

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. DOI: https://doi.org/10.1023/B:JADD.0000022607.19833.00

Barr, D. J., Levy, R., Scheepers, C., & **Tily, H. J.** (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. DOI: https://doi.org/10.1016/j.jml.2012.11.001

Bartels, C. (1999). *The intonation of English statements and questions.* Garland, New York. DOI: https://doi.org/10.4324/9781315053332

Beckman, M. E., & **Ayers, G.** (1997). Guidelines for ToBI labelling. *The OSU Research Foundation, 3*, 30.

Best, C. T., & **Tyler, M. D.** (2007). Nonnative and second-language speech perception. *Language Learning & Language Teaching* (pp. 13–34). DOI: https://doi.org/10.1075/lllt.17.07bes

Beyssade, C., & **Marandin, J. M.** (2006). The speech act assignment problem revisited: Disentangling speaker's commitment from speaker's call on addressee. *Empirical issues in syntax and semantics, 6*(37–68).

Bishop, J. (2016). Individual differences in top-down and bottom-up prominence perception. *Speech Prosody 2016* (pp. 668–672). DOI: https://doi.org/10.21437/SpeechProsody.2016-137

Blum-Kulka, S., House, J., & **Kasper, G.** (1989). Investigating cross cultural pragmatics: An introductory overview. In S. Blum-Kulka, J. House & G. Kasper (Eds.), *Cross-Cultural Pragmatics: Requests and Apologies* (pp. 1–34). Norwood, NJ: Ablex.

Boersma, P., & **Weenink, D.** (2016). Praat: Doing phonetics by computer (Version 6.0. 14). Retrieved from (last access: 29.04. 2018).

Bolinger, D. (1989). *Intonation and its uses: Melody in grammar and discourse.* Stanford University Press.

Borràs-Comes, J., Kiagia, E., & **Prieto, P.** (2019). Epistemic intonation and epistemic gesture are mutually co-expressive: Empirical results from two intonation-gesture matching tasks. *Journal of Pragmatics, 150,* 39–52. DOI: https://doi.org/10.1016/j.pragma.2019.07.004

Borràs-Comes, J., Vanrell, M. del M., & **Prieto, P.** (2014). The role of pitch range in establishing intonational contrasts. *Journal of the International Phonetic Association, 44*(1), 1–20. DOI: https://doi.org/10.1017/S0025100313000303

Brown, L., Winter, B., Idemaru, K., & **Grawunder, S.** (2014). Phonetics and politeness: Perceiving Korean Honorific and non-honorific speech through phonetic cues. *Journal of Pragmatics, 66,* 45–60. DOI: https://doi.org/10.1016/j.pragma.2014.02.011

Büring, D., & **Gunlogson, C.** (2000). Aren't Positive and Negative Polar Questions the Same?. Manuscript, UCSC/UCLA.

Cangemi, F., & **D'Imperio, M.** (2015). Sentence modality and tempo in Neapolitan Italian. In J. Romero & M. Riera (Eds.), *The phonetics-phonology interface: Representations and methodologies* (pp. 109–124). Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/cilt.335.06can

Cangemi, F., Krüger, M., & **Grice, M.** (2015). Listener-specific perception of speaker-specific production in intonation. In Susanne Fuchs, Daniel Pape, Caterina Petrone, & Pascal Perrier (Eds.), *Individual Differences in Speech Production and Perception,* (pp. 123–145). Frankfurt am Main: Peter Lang.

Cason, N., Marmursztejn, M., D'Imperio, M., & **Schön, D.** (2019). Language and Speech Rhythmic Abilities Correlate with L2 Prosody Imitation Abilities in Typologically

Different Languages. *Language and Speech*, *63*(1), 149–165. DOI: https://doi.org/10.1177/0023830919826334

Chen, A., Gussenhoven, C., & Rietveld, T. (2004). Language-specificity in the perception of paralinguistic intonational meaning. *Language and Speech*, *47*(4), 311–349. DOI: https://doi.org/10.1177/00238309040470040101

Chen, H., Rattanasone, X., Cox, F., & Demuth, K. (2017). Effect of early dialectal exposure on adult perception of phonemic vowel length. *The Journal of the Acoustical Society of America*, *142*(3), 1707–1716. DOI: https://doi.org/10.1121/1.4995994

Chikulaeva, A., & D'Imperio, M. (2018). The expression of politeness and pitch height in Russian imperatives. *Speech Prosody 2018* (pp. 438–442). DOI: https://doi.org/10.21437/SpeechProsody.2018-89

Clopper, C. G. (2014). Sound change in the individual: Effects of exposure on cross-dialect speech processing. *Laboratory Phonology*, *5*(1), 69–90. DOI: https://doi.org/10.1515/lp-2014-0004

Clopper, C. G., & Smiljanic, R. (2011). Effects of gender and regional dialect on prosodic patterns in American English. *Journal of phonetics*, *39*(2), 237–245. DOI: https://doi.org/10.1016/j.wocn.2011.02.006

Colon, J., & Bishop, J. (2015). Autistic traits predict prosody perception in neurotypical adults. *Presented at the 27th Meeting of the Association for Psychological Science*, New York City.

Crocco, C., & Badan, L. (2016). 'L'hai messo dove il focus?' Un'analisi prosodica delle domande eco wh. In R. Savy & I. Alfano (Eds.), *La fonetica nell'apprendimento delle lingue [Phonetics and language learning]*. (pp. 191–207). Milano, Officina 21. DOI: https://doi.org/10.17469/O2102AISV000011

Cruttenden, A. (2008). *Gimson's pronunciation of English*. London: Edward Arnold.

D'Imperio, M. (2000). *The role of perception in defining tonal targets and their alignment*. Doctoral dissertation, The Ohio State University.

D'Imperio, M. (2002). Italian intonation: An overview and some questions. *Probus*, *14*(1), 37–69. DOI: https://doi.org/10.1515/prbs.2002.005

D'Imperio, M., Baltazani, M., Gili Fivela, B., Post, B., & Vella, A. (In press). Prosodic systems: Southern Europe, France, Italy, Malta, Albania, Greece (excl. Basque, Iberia, the Balkans) (Chapter 17). In C. Gussenhoven & A. Chen (Eds.), *Oxford Handbook on Prosody*, Oxford University Press.

D'Imperio, M., Cavone, R., & Petrone, C. (2014). Phonetic and phonological imitation of intonation in two varieties of Italian. *Frontiers in psychology*, *5*, 1226. DOI: https://doi.org/10.3389/fpsyg.2014.01226

Dainora, A. (2001). *An empirically based probabilistic model of intonation in American English*. Doctoral dissertation, University of Chicago.

Dayal, V. (2016). *Questions*. Oxford Surveys in Semantics and Pragmatics, OUP. DOI: https://doi.org/10.1093/acprof:oso/9780199281268.001.0001

Degen, J., & Tanenhaus, M. K. (2015). Availability of alternatives and the processing of scalar implicatures: A visual world eye-tracking study. *Cognitive Science*, *40*(1), 1–30. Advance online publication. DOI: https://doi.org/10.1111/cogs.12227

Dittinger, E., Barbaroux, M., D'Imperio, M., Jäncke, L., Elmer, S., & Besson, M. (2016). Professional music training and novel word learning: From faster semantic encoding to longer-lasting word representations. *Journal of cognitive neuroscience*, *28*(10), 1584–1602. DOI: https://doi.org/10.1162/jocn_a_00997

Dittinger, E., Valizadeh, S. A., Jäncke, L., Besson, M., & Elmer, S. (2018). Increased functional connectivity in the ventral and dorsal streams during retrieval of novel

words in professional musicians. *Human brain mapping, 39*(2), 722–734. DOI: https://doi.org/10.1002/hbm.23877

Drager, K., Hay, J., & **Walker, A.** (2010). Pronounced rivalries: Attitudes and speech production. *Reo, Te,* Vol. 53, 2010: 27–53. Availability: https://search.informit.com.au/documentSummary; dn = 830379345270042;res = IELIND

Enfield, N. J., Stivers, T., & **Levinson, S. C.** (2010). Question response sequences in conversation across ten languages: An introduction. *Journal of Pragmatics, 42*(10), 2615–2619. DOI: https://doi.org/10.1016/j.pragma.2010.04.001

Estève-Gibert, N., Schafer, A., Hemforth, B., Portes, C., Pozniak, C., & **D'Imperio, M.** (2018). Individual empathic skills determine how prosody is used to process semantically ambiguous words. *Oral presentation at Socially Situated Language Processing (SSLP 2018), Pre-AmLaP workshop*, ZAS, Berlin, 4–5 Sept. 2018.

Estève-Gibert, N. J., Schafer, A., Hemforth, B., Portes, C., Pozniak, C., & **D'Imperio, M.** (2020). Empathy influences how listeners interpret intonation and meaning when words are ambiguous. *Memory & Cognition* (pp. 1–15). DOI: https://doi.org/10.3758/s13421-019-00990-w

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research, 92,* 233–277.

German, J. S., Carlson, K., & **Pierrehumbert, J. B.** (2013). Reassignment of consonant allophones in rapid dialect acquisition. *Journal of Phonetics, 41*(3–4), 228–248. DOI: https://doi.org/10.1016/j.wocn.2013.03.001

Gili Fivela, B. (2008). Intonation in production and perception: The case of Pisa Italian. Alessandria: Edizioni dell'Orso.

Gili Fivela, B., Avesani, C., Barone, M., Bocci, G., Crocco, C., D'Imperio, M., Giordano, R., Marotta, G., Savino, M., & **Sorianello, P.** (2015). Intonational phonology of the regional varieties of Italian. In S. Frota & P. Prieto (Eds.), *Intonation in Romance.* (pp. 140–197). Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199685332.003.0005

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological review, 105*(2), 251. DOI: https://doi.org/10.1037/0033-295X.105.2.251

Goldinger, S. D. (2000). The role of perceptual episodes in lexical processing. In *ISCA Tutorial and Research Workshop (ITRW) on Spoken Word Access Processes.*

Goodhue, D., & **Wagner, M.** (2018). Intonation, yes and no. *Glossa: A journal of general linguistics, 3*(1), 5. 1–45. DOI: https://doi.org/10.5334/gjgl.210

Goral, F., & **Schellenberg, J.** (2018). goeveg: Functions for Community Data and Ordinations. *R package version 0.4.2.* https://CRAN.R-project.org/package = goeveg

Grice, M., & **Baumann, S.** (2007). An Introduction to Intonation – Functions and Models. In J. Trouvain & U. Gut (Eds.), *Non-Native Prosody. Phonetic Description and Teaching Practice.* Berlin, New York: De Gruyter (= Trends in Linguistics. Studies and Monographs [TiLSM] 186). 25–51. DOI: https://doi.org/10.1515/9783110198751.1.25

Grice, M., & **Savino, M.** (1997). Can pitch accent type convey information status in yes-no questions? In *Proceedings of the ACL Workshop 'Concept-to-Speech Generation Systems'*, Madrid, 14 July 1997 (pp. 29–38).

Grice, M., D'Imperio, M., Savino, M., & **Avesani, C.** (2005). Strategies for intonation labelling across varieties of Italian. In S. A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing (Vol. 1).* Oxford: Oxford University Press, 362–389. DOI: https://doi.org/10.1093/acprof:oso/9780199249633.003.0013

Gunlogson, C. (2004). *True to form: Rising and falling declaratives as questions in English.* Routledge. DOI: https://doi.org/10.4324/9780203502013

Gunlogson, C. (2008). A question of commitment. *Belgian Journal of Linguistics, 22*(1), 101–136. DOI: https://doi.org/10.1075/bjl.22.06gun

Gussenhoven, C. (2002). Intonation and interpretation: Phonetics and phonology. *Speech Prosody 2002* (pp. 47–57). Aix-en-Provence. https://www.isca-speech.org/archive/sp2002/sp02_047.html

Gussenhoven, C. (2016). Foundations of intonational meaning: Anatomical and physiological factors. *Topics in Cognitive Science, 8*(2), 425–434. DOI: https://doi.org/10.1111/tops.12197

Halle, M. (1985). Speculations about the representation of words in memory. *Phonetic linguistics* (pp. 101–114).

Hay, J., & **Drager, K.** (2010). Stuffed toys and speech perception. *Linguistics, 48*(4), 865–892. DOI: https://doi.org/10.1111/j.1749-818X.2010.00210.x

Heritage, J. (2012). Epistemics in action: Action formation and territories of knowledge. *Research on Language & Social Interaction, 45*(1), 1–29. DOI: https://doi.org/10.1080/08351813.2012.646684

Hirschberg, J., & **Ward, G.** (1992). The influence of pitch range, duration, amplitude and spectral features on the interpretation of the rise-fall-rise intonation contour in English. *Journal of phonetics, 20*(2), 241–251. DOI: https://doi.org/10.1016/S0095-4470(19)30625-4

Holliday, N., & **Villarreal, D.** (2020). Intonational variation and incrementality in listener judgments of ethnicity. Laboratory Phonology: Journal of the Association for Laboratory Phonology, *11*(1), 3. DOI: https://doi.org/10.5334/labphon.229

Hothorn, T., Bretz, F., & **Westfall, P.** (2008). Simultaneous Inference in General Parametric Models. *Biometrical Journal, 50*(3), 346–363. DOI: https://doi.org/10.1002/bimj.200810425

Idemaru, K., Winter, B., Brown, L., & **Oh, G. E.** (2020). Loudness trumps pitch in politeness judgments: Evidence from Korean deferential speech. *Language and Speech, 63*(1), 123–148. DOI: https://doi.org/10.1177/0023830918824344

Johnson, K. (1990). The role of perceived speaker identity in F0 normalization of vowels. *Journal of the Acoustical Society of America, 88*(2), 642–654. DOI: https://doi.org/10.1121/1.399767

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 145–165).

Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics, 34*, 485–499. DOI: https://doi.org/10.1016/j.wocn.2005.08.004

Johnson, K., Strand, E. A., & **D'Imperio, M.** (1999). Auditory–visual integration of talker gender in vowel perception. *Journal of phonetics, 27*(4), 359–384. DOI: https://doi.org/10.1006/jpho.1999.0100

Jun, S. A., & **Bishop, J.** (2015). Priming Implicit Prosody: Prosodic boundaries and individual differences. *Language and Speech, 58*(4), 459–473. DOI: https://doi.org/10.1177/0023830914563368

Kidd, E., Donnelly, S., & **Christiansen, M. H.** (2018). Individual differences in language acquisition and processing. *Trends in Cognitive Sciences, 22*(2), 154–169. DOI: https://doi.org/10.1016/j.tics.2017.11.006

Kleinschmidt, D. F. (2019). Structure in talker variability: How much is there and how much can it help? *Language, cognition and neuroscience, 34*(1), 43–68. DOI: https://doi.org/10.31234/osf.io/a4tkn

Krifka, M. (2017). Negated polarity questions as denegations of assertions. In C. Lee, F. Kiefer & M. Krifka (Eds.), *Contrastiveness in Information Structure, Alternatives and Scalar Implicatures* (pp. 359–398). Heidelberg: Springer. DOI: https://doi.org/10.1007/978-3-319-10106-4_18

Kurumada, C., Brown, M., & **Tanenhaus, M.** (2012). Pragmatic interpretation of contrastive prosody: It looks like speech adaptation. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 34, No. 34).

Ladd, D. R. (1981). A first look at the semantics and pragmatics of negative questions and tag questions. In *Proceedings of the 17th Annual Meeting of Chicago Linguistic Society* (pp. 164–171).

Ladd, D. R. (2008). *Intonational phonology.* Cambridge, England: Cambridge University Press. (Original work published 1996). DOI: https://doi.org/10.1017/CBO9780511808814

Ladd, D. R. (2014). *Simultaneous structure in phonology.* OUP Oxford. DOI: https://doi.org/10.1093/acprof:oso/9780199670970.001.0001

Ladusaw, W. A. (2004). Biased questions. *Talk given at UCSC,* Santa Cruz.

Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., & **David, A. S.** (2004). Measuring empathy: Reliability and validity of the Empathy Quotient. *Psychological Medicine, 34*(5), 911–920. DOI: https://doi.org/10.1017/S0033291703001624

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software, 69*(1). DOI: https://doi.org/10.18637/jss.v069.i01

Levon, E. (2016). Gender, interaction and intonational variation: The discourse functions of High Rising Terminals in London. *Journal of Sociolinguistics, 20*(2), 133–163. DOI: https://doi.org/10.1111/josl.12182

Levy, H., & **Hanulíková, A.** (2019). Variation in children's vowel production: Effects of language exposure and lexical frequency. *Laboratory Phonology, 10*(1), 9. DOI: https://doi.org/10.5334/labphon.131

Liberman, M., & **Pierrehumbert, J.** (1984). Intonational Invariance under Changes in Pitch Range and Length. In M. Aronoff & R. Oehrle (Eds.), *Language Sound Structure* (pp. 157–233). Cambridge MA: MIT Press.

Liberman, M., & **Sag, I.** (1974). Prosodic form and discourse function. *Chicago Linguistics Society, 10,* 416–427.

Marwick, B., & **Krishnamoorthy, K.** (2019). cvequality: Tests for the Equality of Coefficients of Variation from Multiple Groups. *R package version 0.2.0.* https://CRAN.R-project.org/package=cvequality

Mathôt, S., Schreij, D., & **Theeuwes, J.** (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior research methods, 44*(2), 314–324. DOI: https://doi.org/10.3758/s13428-011-0168-7

Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics, 32,* 543–563. DOI: https://doi.org/10.1016/j.wocn.2004.02.002

Mennen, I. (2015). Beyond segments: Towards an L2 intonation learning theory (LILt). In E. Delais-Roussarie, M. Avanzi & S. Herment (Eds.), *Prosody and Languages in Contact: L2 acquisition, attrition, languages in multilingual situations* (pp. 171–188). Springer Verlag. DOI: https://doi.org/10.1007/978-3-662-45168-7_9

Michelas, A., Portes, C., & **Champagne-Lavau, M.** (2016). When pitch accents encode speaker commitment: Evidence from French intonation. *Language and speech, 59*(2), 266–293. DOI: https://doi.org/10.1177/0023830915587337

Ohala, J. J. (1983). Cross-language use of pitch: An ethological view, *Phonetica*, *40*, 1–18. DOI: https://doi.org/10.1159/000261678

Orrico, R., Cataldo, V., & **D'Imperio, M.** (to appear). The effect of early dialect exposure in Salerno Italian question intonation. *Proceedings of the XVI AISV Conference*, Rende (CS).

Orrico, R., & **D'Imperio, M.** (2020). Tonal specification of speaker commitment in Salerno Italian wh-questions. *Speech Prosody, 2020* (pp. 361–365). DOI: https://doi.org/10.21437/SpeechProsody.2020-74

Orrico, R., Savy, R., & **D'Imperio, M.** (2019a). Salerno Italian: Intonational phonology and dimensions of variation. In D. Piccardi, F. Ardolino & S. Calamai (Eds.), *Gli archivi sonori al crocevia tra scienze fonetiche, informatica umanistica e patrimonio digitale [Audio archives at the crossroads of speech sciences, digital humanities and digital heritage]* Studi AISV 6. DOI: https://doi.org/10.17469/O2106AISV000018

Orrico, R., Savy, R., & **D'Imperio, M.** (2019b). The perception of speaker certainty in Salerno Italian intonation. In S. Calhoun, P. Escudero, M. Tabain & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences*, (pp. 2946–2950). Melbourne, Australia 2019.

Oshima, D. Y. (2017). Remarks on epistemically biased questions. In *Proceedings of the 31st Pacific Asia Conference on Language, Information and Computation* (pp. 169–177).

Palmer, F. R. (2001). *Mood and Modality*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781139167178

Perry, T. L., Ohde, R. N., & **Ashmead, D. H.** (2001). The acoustic bases for gender identification from children's voices. *The Journal of the Acoustical Society of America*, *109*(6), 2988–2998. DOI: https://doi.org/10.1121/1.1370525

Petrone, C. (2008). *Le rôle de la variabilité phonétique dans la représentation des contours intonatifs et de leur sens*. France: Doktorarbeit, Université de Provence.

Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation*. PhD thesis, MIT. Distributed 1988, Indiana University Linguistics Club.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. Frequency and the emergence of linguistic structure. *Typological studies in language*, *45*, 137–158. DOI: https://doi.org/10.1075/tsl.45.08pie

Pierrehumbert, J. (2002). Word-specific phonetics. *Laboratory phonology*, *7*. DOI: https://doi.org/10.1515/9783110197105.101

Pierrehumbert, J., Bent, T., Munson, B., Bradlow, A., & **Bailey, J. M.** (2004). The influence of sexual orientation on vowel production. *The Journal of the Acoustical Society of America, 116*(4), 1905–1908. DOI: https://doi.org/10.1121/1.1788729

Pierrehumbert, J., & **Beckman, M.** (1988). Japanese Tone Structure. *Linguistic Inquiry Monograph 15*. Cambridge: MIT Press.

Pierrehumbert, J., & **Hirschberg, J.** (1990). The meaning of intonational contours in the interpretation of discourse. In *Intentions in communication* (pp. 271–311).

Portes, C., & **Beyssade, C.** (2015). Is intonational meaning compositional? *Verbum* XXXVII, 2015, n°2, 207–233.

Portes, C., & **German, J. S.** (2019). Implicit effects of regional cues on the interpretation of intonation by Corsican French listeners. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *10*(1), 22. 1–26. DOI: https://doi.org/10.5334/labphon.162

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. DOI: https://doi.org/10.1016/j.pragma.2014.08.013

Prieto, P. (2015). Intonational meaning. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*, 371–381. DOI: https://doi.org/10.1002/wcs.1352

Prieto, P., & **Borràs-Comes, J.** (2018). Question intonation contours as dynamic epistemic operators. *Natural Language & Linguistic Theory*, *36*(2), 563–586. DOI: https://doi.org/10.1007/s11049-017-9382-z

R Core Team. (2019). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. URL https://www.R-project.org/

Rossi, G. (2015). Other-initiated repair in Italian. *Open Linguistics*, *1*, 256–282. DOI: https://doi.org/10.1515/opli-2015-0002

Safarová, M. (2006). *Rises and Falls Studies in the Semantics and Pragmatics of Intonation*. PhD thesis, Universiteit van Amsterdam, Institute for Logic, Language and Computation.

Savino, M., & **Grice, M.** (2011). The perception of negative bias in Bari Italian questions. *Studies in Natural Language and Linguistic Theory* (pp. 187–206). DOI: https://doi.org/10.1007/978-94-007-0137-3_8

Singmann, H., Bolker, B., Westfall, J., Aust, F., & **Ben-Shachar, M.** (2019). afex: Analysis of Factorial Experiments. *R package version 0.24-1*. https://CRAN.R-project.org/package=afex

Sobrero, A. (1993). *Introduzione all'italiano contemporaneo. Le strutture*. Roma-Bari, Laterza.

Steedman, M. (2007). Information-Structural Semantics for English Intonation. In C. Lee, M. Gordon & D. Büring (Eds.), *Topic and Focus: Cross-Linguistic Perspectives on Meaning and Intonation* (pp. 245–264). Dordrecht: Springer. DOI: https://doi.org/10.1007/978-1-4020-4796-1_13

Sudo, Y. (2013). Biased polar questions in English and Japanese. *Beyond expressives: Explorations in use-conditional meaning* (pp. 275–295). DOI: https://doi.org/10.1163/9789004183988_009

Tanenhaus, M. K., Kurumada, C., & **Brown, M.** (2015). Prosody and intention recognition. In *Explicit and implicit prosody in sentence processing* (pp. 99–118). Cham: Springer. DOI: https://doi.org/10.1007/978-3-319-12961-7_6

Todd, S., Pierrehumbert, J. B., & **Hay, J. B.** (2019). Word frequency effects in sound change as a consequence of perceptual asymmetries: An exemplar-based model. *Cognition*, *185*, 1–20. DOI: https://doi.org/10.1016/j.cognition.2019.01.004

Vanrell, M. D. M. (2011). *The phonological relevance of tonal scaling in the intonational grammar of Catalan*. Unpublished doctoral dissertation, Universitat Autònoma de Barcelona.

Vanrell, M. M., Mascaro, I., Torres-Tamarit, F., & **Prieto, P.** (2013). Intonation as an encoder of speaker's certainty: Information and confirmation yes-no questions in Catalan. *Language and Speech*, *56*(2), 163e190. DOI: https://doi.org/10.1177/0023830912443942

Walker, M., Szakay, A., & **Cox, F.** (2019). Can kiwis and koalas as cultural primes induce perceptual bias in Australian English speaking listeners? *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *10*(1). DOI: https://doi.org/10.5334/labphon.90

Ward, G., & **Hirschberg J.** (1985). Implicating Uncertainty: The Pragmatics of Fall-Rise Intonation. *Language*, *64*(1), 747–776. DOI: https://doi.org/10.2307/414489

Ward, G., & **Hirschberg, J.** (1986). Reconciling Uncertainty with Incredulity: A Unified Account of the L*+H L H% Intonational Contour. *Paper presented at the Annual Meeting of the Linguistic Society of America*.

Warren, P. (2017). The interpretation of prosodic variability in the context of accompanying sociophonetic cues. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *8*(1). DOI: https://doi.org/10.5334/labphon.92

Winter, B., & **Grawunder, S.** (2012). The phonetic profile of Korean formality. *Journal of Phonetics*, *40*, 808–815. DOI: https://doi.org/10.1016/j.wocn.2012.08.006

Yu, A., & **Zellou, G.** (2019). Individual Differences in Language Processing: Phonology. *Annual Review of Linguistics 5*(1), 131–150. DOI: https://doi.org/10.1146/annurev-linguistics-011516-033815