

THE PERCEPTION OF SPEAKER CERTAINTY IN SALERNO ITALIAN INTONATION

Riccardo Orrico^{1,2}, Renata Savy¹, Mariapaola D'Imperio^{2,3}

¹University of Salerno, Italy

²Aix Marseille University, CNRS, LPL UMR 7309, Aix en Provence, France

³Rutgers University, New Brunswick (NJ), USA

rorrico@unisa.it, rsavy@unisa.it, mdimperio2@gmail.com

ABSTRACT

One of the main uses of intonation form is to signal pragmatic information. Recent research on intonational meaning has shown that epistemic information about propositional content of utterances within a discourse can also be tonally encoded. The purpose of this study is to investigate whether speaker certainty about the answer to a polar question can be recovered by listeners of Salerno Italian and whether identification is affected by tune type. In an online perception survey, three similarly frequent polar question tunes were rated according to degree of perceived speaker certainty relative to the question response. Results support the hypothesis that tune type affects degree of certainty perception. The study also supports the hypothesis that additional factors – both sociophonetic (i.e., living abroad) and idiosyncratic – might affect variability in intonation perception within a language community.

Keywords: Question intonation, Salerno Italian, epistemic bias, individual variability, perception.

1. INTRODUCTION

It is widely acknowledged that intonation is an essential element for communication and an integral part of language grammar. However, several issues regarding its role in the expression of linguistic functions are still under debate. A hotly debated issue is the stability of the mapping between intonation form and pragmatic meaning.

Most of the research on intonational meaning couched within the Autosegmental-Metrical framework builds on [18], which assigns a specific abstract meaning to single tones, and with the intonation to meaning mapping being stable across contexts and speakers. Moreover, in this approach, the pragmatic meaning of the utterance as a whole builds up in a compositional way, by adding the meanings of each tonal constituent. [18] has also highlighted the property of tones to express the relationship between the propositional content of an utterance and mutual beliefs of discourse participants,

paving the way for investigating the dialogical status of intonation and the tonal signalling of speaker and listener epistemic disposition towards the proposition expressed in an utterance. Specifically, the type of epistemic disposition explored in the present study is the degree of speaker commitment towards the proposition expressed in a yes/no question, that is the degree of certainty, or epistemic bias, with respect to the expected answer to the question.

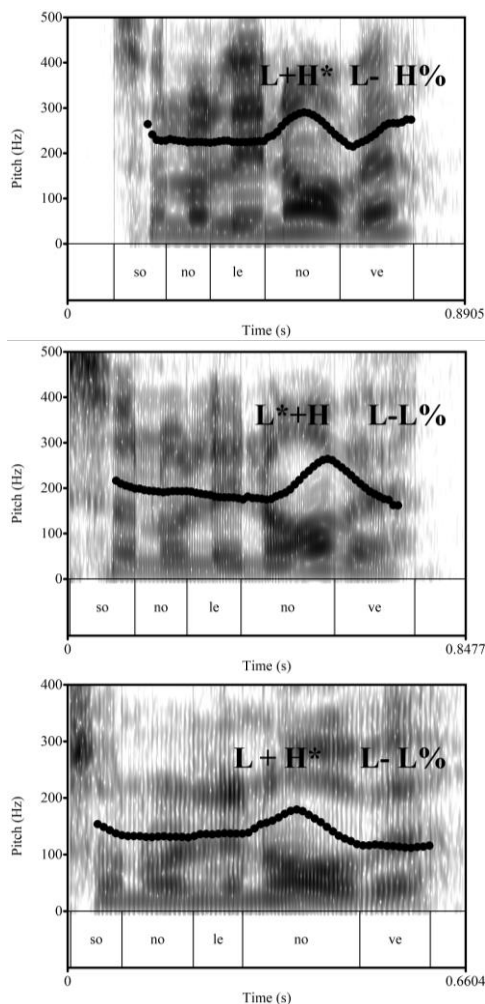
The link between intonation and epistemic commitment has been recently shown for different languages. As for English, [2] provides a compositional account of statements and questions, linking the use of a L- phrase accent to the expression of assertiveness of the proposition expressed. On the other hand, [12] provides a semantic account of rising and falling declaratives, linking the tonal configuration to commitment to the propositional content either on the part of the speaker or of the hearer. As for Romance languages, evidence has been provided for Catalan [19], in which question intonation appears to be linked to both speaker commitment and agreement among participants. Similarly, for Bari Italian [21], pitch accent choice has been argued to be crucial for differentiating neutral information-seeking questions from biased questions.

However, recent research has also shown that individual variability needs to be taken into account when dealing with the relationship between tones/tunes and their meaning within a language community. For example, [8] shows that the mapping between intonational form and pragmatic function is far from being stable, given the absence of a one-to-one relationship. Additionally, [11] reports the presence of speaker-specific strategies in the tonal encoding of informative foci in German.

The present study reports the results of a perception survey testing whether degree of speaker certainty in answering yes/no questions can be perceptually recovered from intonational cues in Salerno Italian (SI). Differently from other Italian regional varieties, the intonational system of SI has only recently been investigated. An account of the

intonation of SI polar questions is reported in [17], where three main phonological tunes have been attested. On the one side, we found two rise-falls, *i.e.* L*+H L-L% and L+H* L-L%, whose phonological contrast depends on the alignment of the H target within the nuclear pitch accent (one early and the other late with respect to vowel onset), while having the same falling boundary; on the other, we also found a rise-fall-rise, *i.e.* L+H* L-H%, presenting an early peak accent and a rising boundary tone (Fig. 1). Crucially, the distribution of these tunes reported in our previous study appears to be also dependent on individual speaker variability rather than on mere pragmatic function.

Figure 1: F0 contours for the yes/no question *Sono le nove?* “Is it nine o’clock?” uttered with a L*+H L-L% (top), L+H* L-H% (middle), or a L+H* L-L% (bottom) tune.



Variability of phonetic categories at the individual level raises the question of how listeners manage to decode a potentially ambiguous message. Recent work has proposed a model for quantifying how listeners rely on the covariation of phonetic cues and socio-indexical features (age, gender, and dialect) of

the speakers to cope with the uncertainty deriving from the variability of the speech signal, therefore, by assuming that phonetic variation is systematic rather than random [13]. The effects that social features have on phonetic production might be mirrored in perception and shape the way listeners make inferences about speakers’ intentions. Here we employed a perception survey to test the effect of exposure to other varieties of Italian or languages, hypothesizing that phonological traces from past experience with other systems might affect the perception of the native variety, as predicted by exemplar-based models [10] and supported by imitation studies [5, 7].

The specific hypotheses we test are: 1) degree of speaker certainty about the answer is conveyed by question tune choice, and 2) perception of speaker certainty relative to a particular tune is not stable, but depends on listeners’ prior exposure to other phonological systems [7, 16]. Since Italian varieties are very different from each other in their intonational phonology [9], we tested the effect of having lived outside of Salerno for at least one year.

2. METHOD

2.1. Experimental stimuli

The experimental stimuli were produced by a female native speaker of SI. 24 target items were recorded using the 3 question tunes described above, *i.e.* L+H* L-H%, L*+H L-L%, and L+H* L-L%, amounting to a total of 72 stimuli. Each item was composed of a verb and a trisyllabic object, with the nuclear pitch accent being always aligned with the antepenultimate syllable of the object (*Stai cercando le MANdorle?*, ‘Are [you] looking for the almonds?’). The 72 stimuli were distributed in 3 lists, each containing 24 target stimuli. Each tune appeared 8 times in each list while each item appeared only once in the same list.

2.2. Participants

We recruited 45 listeners (23 males), aged from 21 to 71 (median: 27). All participants were native speakers of SI and all of them were living in Salerno at the time of the experiment.

2.3. Experimental procedure

The experiment was submitted to the listeners through an online research platform (Gorilla.sc). After the explanation and demonstration of the task, each listener was randomly assigned to one of the three lists of stimuli. The listeners were to judge the stimuli according to perceived speaker certainty about the answer to the question utterance, using a

slider ranging from 0 to 100 (0 = ‘She expects no’; 100 = ‘She expects yes’).

Participants were also asked to fill a sociolinguistic questionnaire. Among the items, we asked whether they had lived outside of Salerno for a period of 12 months or longer.

2.4. Statistical analysis

Speaker certainty (100-point scale, treated as continuous) was analysed using mixed-effect regression models in R [20], using the packages *lme4* [3] and *afex* [22]. The model included Tune type (3 levels), List (3 levels), and Other cities (having lived or not in another city, 2 levels) as fixed effects. We followed the suggestions of [1] and first tested the maximal random effect structure allowed by the design by including Listeners and Items as random intercepts and slopes for the effect of Tune type. Given that the maximal model did not converge, we then suppressed correlations among random parameters¹.

General effects were obtained using the function *mixed* from the package *afex*, using the Kenward-Roger approximation for fixed-effects testing.

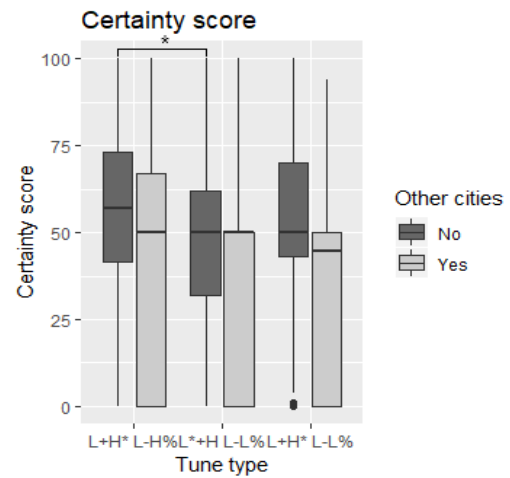
3. RESULTS

Results are shown in Fig. 2. First, the model revealed a significant main effect of Tune type, $F(2, 32.07) = 4.87, p < .01$. Further inspection of the differences among levels of this factor were conducted through pairwise comparisons adopting the Bonferroni correction for p -values, using the package *lsmeans* [14]. Overall, L^*+H $L-L\%$ received lower certainty values, however the only pair of tunes that yielded significantly different results was L^*+H $L-L\%$ vs. $L+H^*$ $L-H\%$ (Est. = 6.23, SE = 1.96, $t = 3.16, p < .01$), while the inspection of the other two contrasts revealed no significant difference.

The model also revealed a significant main effect of Other cities, $F(1, 47.32) = 5.89, p < .02$. No statistical significance was obtained, on the other hand, for the interaction between Tune type and Other cities, nor for the main effect of List.

Note also that listeners who lived in other cities globally rated all stimuli with a lower certainty value of about 14 points relative to all other listeners, regardless of Tune type. The responses within this group were also more variable, as it can be noticed from the size of the boxes in Fig. 2. We hence tested equality of the coefficients of variation between the two groups through the Asymptotic test and the Modified signed-likelihood ratio test (M-SLRT), both

Figure 2: Speaker Certainty scores by Tune Type and Other Cities.



available in the package *cvequality* [15]. The summary of the tests is reported in Table 1, showing a significant difference in variation across the two conditions.

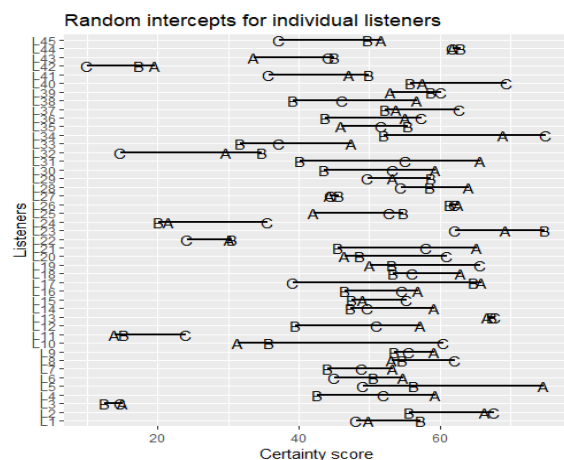
Table 1: Summary of tests on the equality of coefficients of variation in Certainty score by Other cities

Test name	Test statistics	p-value
Asymptotic	110.6385	<.001
M-SLRT	77.1975	<.001

3.1. Random intercepts for individual listeners

Fig. 3 reports the plot of intercept values for individual listeners.

Figure 3: Intercept values of Speaker Certainty scores for each individual listener. A: $L+H^*$ $L-H\%$; B: L^*+H $L-L\%$; C: $L+H^*$ $L-L\%$.



¹Retained model: Certainty score ~ Tune * Other cities * List + (Tune || Participants) + (Tune || Item).

Note that, despite the global effect of Tune type reached significance, listeners behaved quite differently from each other. In particular, only 31.1% of our listeners followed the exact trend registered in the model (Certainty score: $L^*+H L-L\% < L+H^* L-L\% < L+H^* L-H\%$). Specifically, 57.7% of the listeners assigned a higher Certainty value to early peak accent tunes, while more than 40% of listeners assigned a lower value to the early peak pitch accent tunes.

4. DISCUSSION

In this study, we tested the effect of 3 Salerno Italian yes/no question tune types, *i.e.* $L^*+H L-L\%$, $L+H^* L-L\%$, and $L+H^* L-H\%$ on perceived speaker's epistemic stance, *i.e.* degree of certainty about the answer to a yes/no question. Our results suggest that tune variability systematically affects listener's judgments. In particular, the late peak accent tune with a low boundary scored significantly lower than the early peak pitch accent tune with a high boundary, while the early peak accent tune with a low boundary received an intermediate score. Our findings are consistent with those reported for Bari Italian, in which early peak pitch accents correlate with positively biased yes/no questions [21]. These results point to the fact that, despite the great variability observed in intonation across Italian varieties, also reported by [9], inter-variety commonalities at the pragmatic level can still be found. Nevertheless, differently from Bari Italian where pitch accent alone is argued to signal speakers' bias, in SI an early peak pitch accent signals speaker certainty only if in combination with rising boundary tones. Indeed, our results seem to point towards the fact that it is the whole tune, rather than the single tone, that is employed to extract epistemic bias. Further evidence, however, is needed to better address this question.

Additionally, also exposure to other phonological systems seems to have an effect: listeners who lived in other cities for a period longer than one year assigned lower certainty scores to all tunes while also showing greater variability in the responses. This seems to point to a permeability in the intonation grammar that might lead to retuning of categories, as found in imitation and selective perception studies [5, 7, 24].

These results are also in line with what has been reported in second language learning studies. Specifically, the idea that the ability to reshape one's phonological system is not lost in adulthood is a fundamental claim behind some of the L2 learning models, such as the Speech Learning Model [6] and the Perceptual Assimilation Model for L2 [4]. In particular, [4] predicts that listeners' perception of

phonetic information is a dynamical system and the contact with other languages or even other dialects of the same language can lead to perceptual changes. Also, perceptual changes usually appear very fast, after only 6-12 months of exposure to the new system, which is about the same minimal amount of time that our listeners spent living in other cities.

Despite [4]'s claims refer to segmental phonetics, we have no reasons to believe that the same might not apply to intonation categories. Firstly, some of the assumptions in [4], such as the assimilation of a non-native category to a native one, have been successfully tested also for lexical tones, showing that the mechanisms involved in the perception of non-native segments are also exploited for suprasegmentals [23]. In addition, evidence for the effects of a second language on first language intonation production are provided by [16]'s study on tonal alignment in Dutch non-native speakers of Greek, in which L2 experience leads to bidirectional influences on tonal alignment values, *i.e.* modifications in the way tonal categories are realized in L1 and not simply in L2.

Finally, our results point to differences in the behaviour of individual listeners within the general population, which are harder to account for. We can speculate that most of the individual differences observed in our study might depend on cognitive and other sociophonetic aspects on which future work should be conducted.

5. CONCLUSION

A perception study tested the relative impact of tune type on perceived epistemic stance (speaker certainty relative to the answer to a yes/no question) and its relation to both sociophonetic and individual variability in SI. Results revealed that different tunes are systematically linked to different degrees of perceived speaker certainty, despite a great amount of individual variability, whose sources still remain unexplained and should be investigated in cognitive models.

In addition, an effect of sociolinguistic variables was registered. In particular, we report the effect of group differences related to contact with other varieties, given that one of the two listeners groups included speakers having lived in other cities than Salerno for more than a year. Although further research is needed to better account for the contribution of intonation in expressing epistemicity in SI, our findings clearly underline the role of both sociophonetic and idiosyncratic factors. This is in line with the idea that a direct link between intonational form and pragmatic/epistemic meaning cannot be uniformly modelled within a language community.

7. REFERENCES

- [1] 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.
- [2] Bartels, C. 2014. *The intonation of English statements and questions: A compositional interpretation*. Routledge.
- [3] Bates, D., Maechler, M., Bolker, B., Walker, S. 2014. lme4: Linear mixed-effects models using Eigen and S4. R package version, 1(7), 1-23.
- [4] Best, C. T., Tyler, M. D. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In: Bohn, O. S., Munro, M. J. (eds) *Language experience in second language speech learning: In honor of James Emil Flege*, 1334, 1-47.
- [5] D'Imperio, M., Cavone, R., Petrone, C. 2014. Phonetic and phonological imitation of intonation in two varieties of Italian. *Frontiers in psychology*, 5, 1226.
- [6] Flege, J. E. 1995. Second language speech learning: Theory, findings, and problems. In W. Strange (eds), *Speech perception and linguistic experience: Issues in cross-language research*. Timonium, MD: York Press, 233-276.
- [7] German, J., D'Imperio, M. 2016. Like a square peg in a round hole: Why contour shape matters for learning new intonation patterns. Poster presented at LabPhon 15, July 13-16 2016, Cornell University, USA.
- [8] Gili Fivela, B. 2008. *Intonation in production and perception: The case of Pisa Italian*. Edizioni dell'Orso.
- [9] 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: Frota, S., Prieto, P. (eds) *Intonation in Romance*. Oxford University Press, 140-197.
- [10] Goldinger, S. D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological review*, 105(2), 251.
- [11] Grice, M., Ritter, S., Niemann, H., Roettger, T. B. 2017. Integrating the discreteness and continuity of intonational categories. *Journal of Phonetics*, 64, 90-107.
- [12] Gunlogson, C. 2004. *True to form: Rising and falling declaratives as questions in English*. Routledge.
- [13] 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.
- [14] Lenth, R.V. 2016. Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, 69(1), 1-33. doi:10.18637/jss.v069.i01
- [15] Marwick, B., Krishnamoorthy, K. 2018. cvequality: Tests for the Equality of Coefficients of Variation from Multiple Groups. R package version 0.1.3. <https://github.com/benmarwick/cvequality>.
- [16] Mennen, I. 2004. Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of phonetics*, 32(4), 543-563.
- [17] Orrico, R., Savy, R., D'Imperio, M. *in press*. Salerno Italian: intonational phonology and dimensions of variation. *Studi AISV* 4.
- [18] Pierrehumbert, J., Hirschberg, J. B. 1990. The meaning of intonational contours in the interpretation of discourse. *Intentions in communication*, 271-311
- [19] Prieto, P., Borràs-Comes, J. 2018. Question intonation contours as dynamic epistemic operators. *Natural Language and Linguistic Theory*, 36 (2), 563-586.
- [20] R Core Team 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- [21] Savino, M., Grice, M. 2011. The perception of negative bias in Bari Italian questions. In Frota, S. Prieto, P., Elordieta, G. (eds), *Prosodic categories: Production, perception and comprehension*. Springer: Dordrecht, 187-206.
- [22] Singmann, H., Bolker, B. Westfall, J. & Aust, F. 2018. afex: Analysis of Factorial Experiments. R package version 0.22-1. <https://CRAN.R-project.org/package=afex>
- [23] So, C. K., Best, C. T. 2014. Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories. *Studies in Second Language Acquisition*, 36(2), 195-221.
- [24] Tanenhaus, M. K., Kurumada, C., Brown, M. 2015. Prosody and intention recognition. In: Frazier, L., Gibson, E. (eds) *Explicit and implicit prosody in sentence processing*. Springer, Cham, 99-118.