

The effects of regional Italian prosodic variation on modality identification by L1 English learners

Valentina De Iacovo¹, Paolo Mairano²

¹Laboratorio di Fonetica Sperimentale ‘Arturo Genre’, Università di Torino

²STL - Savoirs, Textes, Langage, Université de Lille

valentina.deiacovo@unito.it, paolo.mairano@univ-lille.fr

Abstract

Yes-no questions in Italian are not marked morpho-syntactically and intonation is the only cue distinguishing declarative vs interrogative modality. However, in different regional varieties of Italian, the intonation patterns of questions vary dramatically and yes-no questions can be realised with final rising or falling contours. In this contribution, we investigate whether adult learners of L2 Italian correctly identify the modality (interrogative vs declarative) of a sentence, when pronounced by native speakers of different regional provenance, using a variety of rising and falling contours. We developed an identification test where participants were exposed to 100 stimuli (10 sentences x 5 varieties x 2 modalities), pronounced by 10 speakers from 5 different regions in Italy. 20 L1 English learners of L2 Italian and 20 L1 Italian control speakers listened to the final syllables of each utterance and identified it as declarative or interrogative. Results show that L1 Italian speakers correctly identify sentence modality at higher rates than learners, and that questions with final falling contours have the lowest correct identification rates for learners. We argue that this may be attributed to L1 transfer (since a rise is the default realisation for yes-no questions in English, even more so without syntactic inversion), as well as to universal patterns.

Index Terms: prosodic variation, regional Italian, perception task, intonational pattern, L2 English learners

1. Introduction

1.1. Yes-no questions in English and Italian

Intonation conveys information at various linguistic and extra-linguistic levels, including modality (e.g., declarative, interrogative, continuative) [1]. While in some languages the interrogative modality is realised via various non-prosodic cues (typically at the morpho-syntactic level, via word order, particles or affixes), other languages rely on intonation alone to express sentence modality. English is an example of the former language type, and uses subject-verb inversion and do constructions to express the interrogative modality. In this type of languages, intonation can accompany non-prosodic cues for expressing modality: traditional descriptions of English attribute falling contours to declarative sentences and rising contours to unmarked yes-no questions [2] [3] [4], with a certain amount of variation, for example more frequent low rises in British English and more frequent high or wide rises in American English (but [4] challenged this view). Other studies observed the use of falling contours for yes-no questions in some cases [5], particularly in the case of “non-genuine” questions [6] [7]. However, falling contours for yes-no questions are less frequent than rises, and the latter are therefore as unmarked. We should also point out that intonation can sometimes be the only cue expressing modality

in English, as in the so-called declarative questions (i.e., without inversion such as *she went there?*), which systematically rise [3]. Additionally, the association between falling and rising contours to statements and questions respectively has also been described as a universal trend among languages (among others [8], [9]). Vice versa, Italian is an example of languages where intonation tends to be the only linguistic cue distinguishing declarative vs interrogative modality. For instance, the sentence *C’è ancora della pasta nella pentola(?)* (There is still pasta in the pot(?)) can be declarative or interrogative, depending on intonation alone. In Standard Italian, a falling contour is canonical for declarative sentences, and a rising contour is canonical for yes-no questions ([10]), following the typological trend observed in many languages and mentioned above. Nevertheless, speech is characterised by variations depending on many factors, including regional variety. Recent studies have looked at intonational variation of Italian varieties (see [11], [12], [13]), and have usually found that regional variation affects more particularly the interrogative modality, where it is possible to recognize specific intonation patterns for specific varieties. For example, some of them make extensive use of falling contours even for questions. In this contribution we shall consider intonational variation across 5 regional varieties of Italian: TO (Turin, north-west), PD (Padova, north-east), PR (Prato, centre), CA (Cagliari, Sardinia) and PA (Palermo, Sicily). These varieties have been chosen in the attempt to cover dialectal areas of Italian across the whole country and because they differ regarding the default final contour for the interrogative modality (see figure 1): falling-rising for PR ([14]) and CA ([15]), rising-falling for TO ([16]), PD ([11]) and PA ([17]).

1.2. Effects of regional variation on L2 phonology

It seems therefore reasonable to believe that the amount of regional variation affecting intonation patterns among Italian speakers can potentially be an issue for learners of Italian as a second language (L2), since they may have had limited exposure to such varieties. In effect, Standard Italian is based on the variety spoken in Florence, Tuscany (central Italy), and this variety is certainly the most frequent in audio materials conceived for L2 learners, as well as in the media. This means that L2 Italian learners may be confronted to difficulties in recognising sentence modality when listening to regional varieties exploiting intonation patterns that differ from Standard Italian, or from the universal trend (rising for questions, falling for declaratives). The effects of regional variation on L2 acquisition have been little explored in the literature, and even less so for prosody. Some studies have concentrated on the effect of L1 regional variation on L2 acquisition, e.g. differences in the perception of L2 Dutch by L1 Peruvian vs L1 Iberian Spanish learners [18], of L2 English by L1 Dutch and L1 Belgian learners [19],

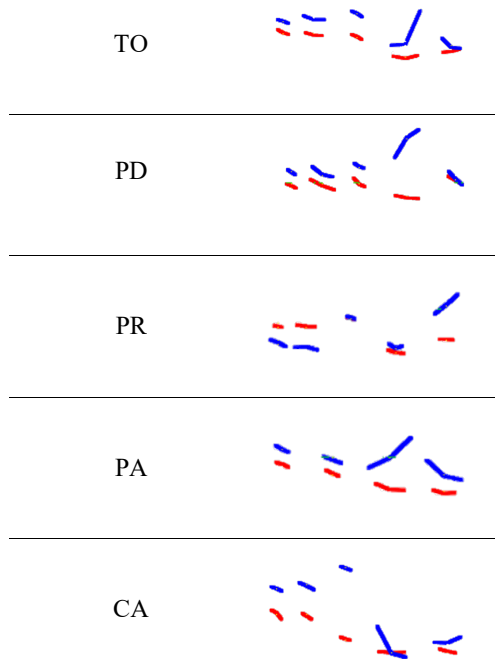


Figure 1: f_0 contours of the 5 last syllables of the sentence "La riunione di martedì è stata confermata.!" for the 5 varieties employed in the task (red=declarative, blue=interrogative modality). The nucleus is on the penultimate syllable.

and differences in production patterns by L2 German learners of L1 Italian with different dialectal backgrounds [20], with some other studies failing to find significant effects (e.g. for Standard vs Cypriot L1 Greek learners of L2 Dutch [21], possibly due to the high acoustic proximity between the vowel systems of these two varieties). But the literature is far scarcer on the effects of L2 regional variation: [22] found that exposure to L2 regional varieties affects how well learners perceive and produce specific segments in an L2, and [23] revealed that less-experienced L2 listeners use L1 vowel categories in the identification of L2 vowels, especially for less familiar varieties. None of these studies examined prosodic patterns, confirming the tendency of L2 phonology acquisition studies to concentrate on segmentals rather than on prosody.

1.3. Goal of our study and expectations

We investigate whether modality identification in L2 Italian by L1 English learners is hindered by L2 regional variation. We developed a listening task with L2 Italian sentences as spoken in various regional varieties. We included five regional varieties of Italian as spoken in the following 5 cities: TO (Turin), PD (Padova), PR (Prato), CA (Cagliari), PA (Palermo). We expect that interrogative utterances produced by speakers of varieties using a rising contour will be more accurately identified as questions, while interrogative utterances produced by speakers of varieties using a falling contour will be more challenging. Such expectations are based on (a) markedness, i.e. the universal trend of questions to rise; (b) exposure, i.e. the general familiarity of L2 learners with the SI intonation, which tends to rise for questions; (c) L1 transfer, given that English yes-no questions are typically associated to rises ([2]).

2. Data and Methodology

In order to test the effect of regional contours on modality recognition in L2 Italian, we developed a modality identification test based on audio recordings. The details of the test and of the participants are given in this section.

2.1. Stimuli

The stimuli consisted of 10 Italian sentences that could work in the declarative or interrogative (yes-no question) modality without any modification other than intonation (e.g. *La riunione di martedì è stata confermata.* "Tuesday's meeting was confirmed(?)"). Each sentence was recorded in the two modalities in the five target regional varieties. For each variety we recorded two speakers, a man and a woman, each of which read 5 sentences in a sound-proof booth at the LFSAG lab in Turin. In total, we therefore recorded 5+5 sentences x 2 modalities x 5 varieties = 100 stimuli. The sentence had variable length, but all recordings were cut in Praat ([24]), so that only the last 5 syllables of each sentence were left (everything else was removed). This was done in order to obtain stimuli of comparable duration, and to remove intonation patterns at the beginning of each sentence that may have carried information about modality. The nucleus is on the penultimate syllable for 8 sentences, and on the last and antepenultimate syllable for the remaining 2 sentences. The audio files were then normalised to an average intensity of 70 dB and converted to mp3 format for use within Gorilla (see below).

2.2. Test format

The Gorilla platform [25] was used to create and host the modality identification task. At the beginning of the test, the instructions included the list of written sentences with the aim of familiarising participants with the stimuli: we did not want them to concentrate on understanding the words, but rather on recognising the modality. After the instructions, the test included a short training phase with 4 trials, and then the 100 real trials. Within each trial, after a fixation of 700 ms, participants listened to an audio stimulus and had to decide if it was a question or a declarative sentence by clicking on a key. The duration of the whole test was approximately 10-15 minutes.

2.3. Participants

We recruited 20 L1 English learners of L2 Italian (10F, 10M, age range: 24-47, henceforth EN group). They were all born in the UK (various cities) and had studied Italian for 1 to 7 years at the time of testing. Additionally, 20 L1 Italian speakers (10F, 10M, age range: 24-42) from different regional backgrounds were recruited as controls (henceforth IT group).

2.4. Pitch contours of the stimuli

In order to investigate the effect of contour on participants' responses, we analysed the intonation patterns of each stimulus. This analysis was carried out auditorily, with the supporting visualisation of the pitch track on Praat (the window was adjusted to 100 – 400 Hz for female speakers and 50 – 250 Hz for male speakers, or as needed to avoid octave jumps and other F0 detection errors). The contour of each nuclear pattern was coded as falling, rising, rise-falling, or fall-rising. Unsurprisingly, declarative sentences generally (though not exclusively) exhibited a falling pattern, with some exceptional fall-rising patterns notably by PD speakers. On the other hand, the interrogative stim-

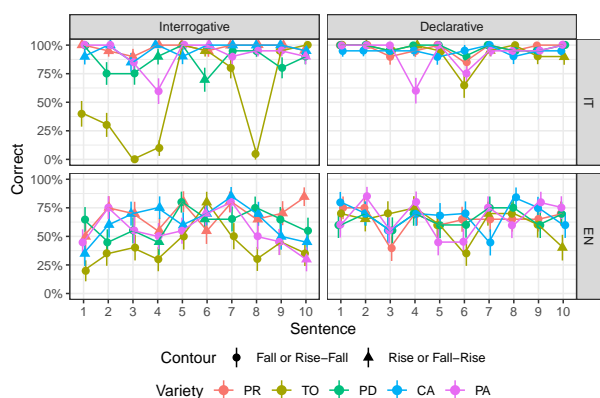


Figure 2: Accuracy by group for declarative and interrogative stimuli. Circles indicate falling (or rise-falling contours), triangles indicate rising (or fall-rising) contours.

uli are characterised by rising or fall-rising patterns, while the rise-falling pattern is especially used by TO speakers, but also in some realisations by PD and PA speakers (see figure 1).

3. Results

3.1. Global results

The data recorded by Gorilla were saved in .csv format and analysed in R [26]. Global accuracy scores for each of the 10 sentences are presented in figure 2 by modality (interrogative vs declarative) and group (L1 English learners and native Italian speakers), also distinguishing rising (or fall-rising) vs falling (or rise-falling) contours. The plots immediately show that accuracy is globally higher for IT than for EN, which is clearly not surprising. Interestingly, the performance of the IT group would be essentially at ceiling level, if it were not for five TO stimuli, which have not been recognised as interrogative (accuracy < 40%, and in three cases even close to 0%). The same five sentences also have particularly low accuracy among the EN group, and in general interrogative stimuli in the TO variety are recognised worse than in the other varieties.

3.2. Effects of L2 regional variation

In order to gain a deeper insight, we analysed the data with mixed-effects models, separately for declarative and interrogative sentences, using lme4 [27], lmerTest [28], and emmeans [29] libraries, using Satterthwaite' approximation to compute degrees of freedom and p values. A first model was built to predict participants' responses for declarative sentences, including variety and group as fixed effects, with random intercepts for participant and sentence, and a by-group random slope for sentence (random slopes for participant were not included to avoid singular fit). Predictions for this model were extracted via ggeffect [30] and are illustrated in figure 3. The model revealed a significant effect of group (IT having significantly higher accuracy than EN, $p < .001$), but no significant effect of the regional variety. This suggests that declarative sentences are recognised at similar accuracy rates across the five varieties considered.

A second analogous model examined participants' responses for interrogative sentences, with variety and group as fixed effects, random intercepts for participant and sentence, and by-group random slope for sentence. Predictions for this

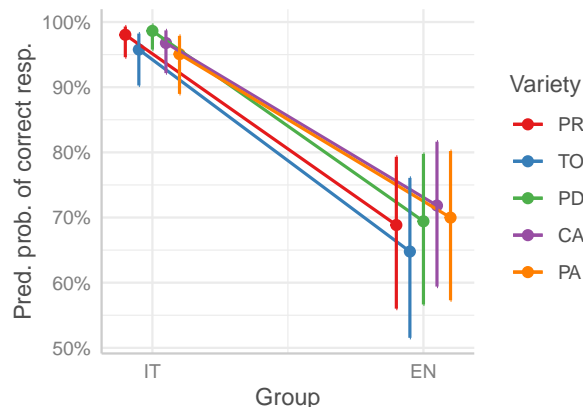


Figure 3: Predicted accuracy for declarative sentences, by group and variety.

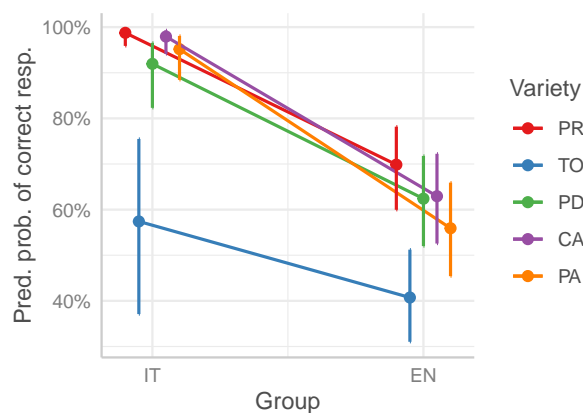


Figure 4: Predicted accuracy for interrogative sentences, by group and variety.

model are illustrated in figure 4. Again, we have a significant and expected effect of group (IT having significantly higher accuracy than EN, $p < .001$), but this time we also have a significant effect of variety, which is evident in the plot, as well as a significant interaction of variety and group. Post-hoc pairwise comparisons with Tukey adjustment revealed that TO stimuli were recognised as interrogative with significantly lower accuracy than the other four varieties by the IT group (all p values < .001) and the EN group (all p values < .03). Additionally, among the IT group accuracy was significantly lower for PD than for PR and CA (both p values < .002), whereas among the EN group it was significantly lower for PA than PR ($p = .05$). This suggests that stimuli uttered with TO intonation patterns were the most difficult to identify as interrogative, both for IT and EN, and that PR (Tuscan) patterns were the easiest for both groups, as can be observed in figure 4.

3.3. Effects of contour

Since TO interrogative sentences are also the ones that most frequently use a falling or rise-falling pattern on the last syllables, these results seem to corroborate the hypothesis that the identification of the interrogative modality is more difficult with non-rising patterns. In order to more directly test this hypothesis, we built a further model to predict responses for interrog-

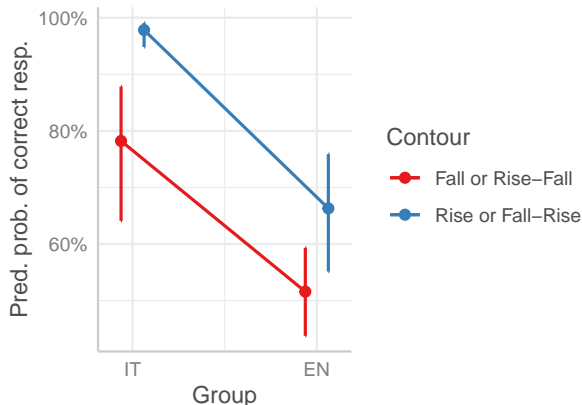


Figure 5: Predicted accuracy for interrogative sentences, by group and contour.

ative sentences: group and contour were included as fixed effects, with random intercepts for participant and sentence; in this case, we were able to include a by-group random slope for sentence, and a by-contour random slope for participant. The predictions of this model are illustrated in figure 5: beyond the significant effect of group ($p < .001$), we observed a significant effect of contour ($p < .001$), and a significant interaction of contour \times group ($p < .001$) suggesting that the effect of contour was stronger among the IT group than among the EN group. Pairwise comparisons revealed that stimuli with a falling contour were correctly identified as interrogative at a much lower rate than stimuli with a rising contour by both the IT group ($p < .001$) and the EN group ($p = .005$).

4. Final discussion

Globally, these results suggest that prosodic patterns due to L2 regional variation can affect the recognition of sentence modality by L2 learners. In our data, TO turns out to be most difficult in this respect, corroborating the idea that patterns that are more marked and/or divergent from the L1 may pose challenges. However, some observations need to be pointed out.

Firstly, among the varieties included in this study, only TO seems to pose relevant challenges for modality recognition for L2 learners. We expected the Tuscan variety (PR) to be the easiest, given that L2 Italian learners are generally exposed to Tuscan Italian in the classroom and in the media (despite potentially different exposures to and preferences for given varieties by each learner). In fact, although accuracy is indeed highest for PR, it does not significantly diverge for PR vs PD and CA, and only slightly for PR vs PA. This may suggest that variation in prosodic detail due to L2 regional varieties may not always constitute a difficulty for learners, as long as they can somehow match it to a prototypical pattern (be it a universal pattern, or a pattern matching L1 characteristics, or a pattern matching their reference L2 variety – which we assume to be Standard Italian). Instead, difficulties arise when the prosodic pattern cannot be matched to this prototype, as is the case for non-rising interrogative intonation for TO. These results go in the same directions as previous studies on modality identification, who found that learners and native speakers tend to associate rising contours to questions ([9], [31]) and who argued that this may be related to a universal trend.

Secondly, the effect of contour is visible not only for EN,

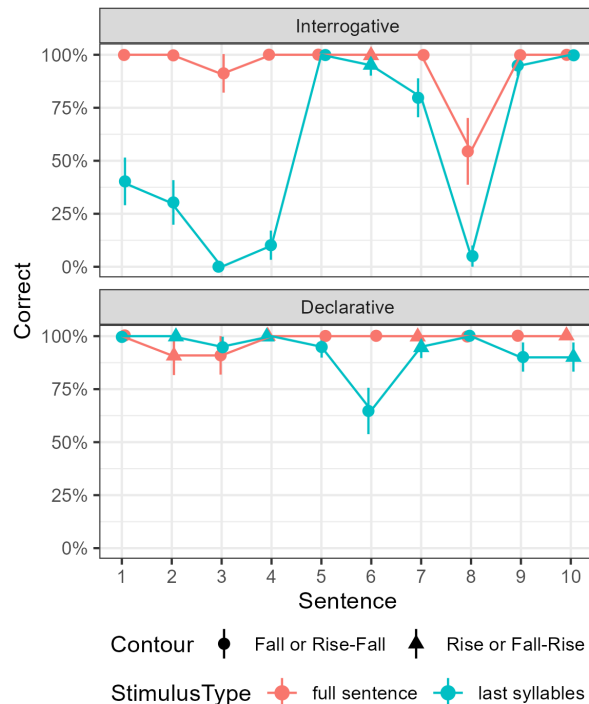


Figure 6: Accuracy by stimulus type (full sentence vs last 5 syllables) for declarative and interrogative stimuli by TO speakers. Circles indicate falling (or rise-falling contours), triangles indicate rising (or fall-rising) contours.

but also for IT participants: TO stimuli with falling contours are difficult to identify as questions not only for learners but also for native speakers. Since the stimuli only included the last five syllables, it may be the case that TO questions are characterised by intonation features elsewhere than in the last five syllables (and therefore not available to participants). In order to address this matter, a further test was run with 10 native Italian speakers, who performed the same test described above, but with TO full sentences rather than the five last syllables. The results of this further test are illustrated in figure 6, and show that native speakers do not seem to have issues in recognising the interrogative modality for TO stimuli if presented with the full sentence, suggesting that Piedmontese regional Italian can express the interrogative modality via prosodic cues at the beginning rather than at the end of the IP – at least in some cases. However, this too is a characteristic which may not be familiar to L2 learners, and which will be explored further in our future studies.

Finally, another interesting question that remains open is whether the disturbing effect of non-rising contours for interrogative stimuli is due to L1 transfer (rise being the standard realisation for yes-no questions in English) or markedness (rises being universally associated to the interrogative modality). Our experiment was not designed to specifically distinguish between these two possible causes, and only future studies will be able to answer this.

5. Acknowledgements

The authors would like to thank the speakers who accepted to be recorded, and all the participants of the perception test.

6. References

- [1] I. Fónagy, "On status and functions of intonation," *Acta Linguistica Hungarica*, vol. 39, no. 1/4, pp. 53–92, 1989.
- [2] A. Cruttenden, "Falls and rises: meanings and universals," *Journal of Linguistics*, vol. 17, no. 1, pp. 77–91, 1981.
- [3] J. C. Wells, *English intonation PB and Audio CD: An introduction*. Cambridge University Press, 2006.
- [4] J. M. Levis, "The intonation and meaning of normal yes/no questions," *World Englishes*, vol. 18, no. 3, pp. 373–380, 1999.
- [5] N. Nakanishi and M. Haneo, "Is rising intonation the default pattern in yes-no questions? the cases of british and american english," *Studies in English Language Teaching*, vol. 44, pp. 61–80, 2021.
- [6] N. Hedberg, J. M. Sosa, and E. Görgülü, "The meaning of intonation in yes-no questions in american english: A corpus study," *Corpus Linguistics and Linguistic Theory*, vol. 13, no. 2, pp. 321–368, 2017.
- [7] A. Banuazizi and C. Creswell, "Is that a real question? final rises, final falls, and discourse function in yes-no question intonation," *CLS*, vol. 35, pp. 1–14, 1999.
- [8] J. J. Ohala, "An ethological perspective on common cross-language utilization of f of voice," *Phonetica*, vol. 41, no. 1, pp. 1–16, 1984.
- [9] C. Gussenhoven and A. Chen, "Universal and language-specific effects in the perception of question intonation," in *6th International Conference on Spoken Language Processing (ICSLP 2000)*, 2000, pp. 91–94.
- [10] E. Magno-Caldogno, F. E. Ferrero, C. Lavagnoli, and K. Vaggés, "F0 contours of statements, yes-no questions, and wh. questions of two regional varieties of italian," *Journal of Italian Linguistics Amsterdam*, vol. 3, no. 1, pp. 57–66, 1978.
- [11] L. Canepari, *L'intonazione: linguistica e paralinguistica*. Liguori, 1985.
- [12] P. Sorianello, *Prosodia: modelli e ricerca empirica: contiene materiali online*. Carocci editore, 2011.
- [13] B. Gili Fivela, C. Avesani, M. Barone, G. Bocci, C. Crocco, M. D'Imperio, R. Giordano, G. Marotta, M. Savino, and P. Sorianello, "Intonational phonology of the regional varieties of italian," in *Intonation in romance*. Oxford University Press, 2015, pp. 140–197.
- [14] V. De Iacovo, *Intonation analysis on some samples of Italian dialects: an instrumental approach*. Edizioni dell'Orso, 2019, vol. 3.
- [15] M. Contini, "L'intonation des phrases affirmatives et interrogatives, avec inversion en sarde. analyse et synthèse," *Bulletin de l'Institut de Phonétique de Grenoble*, vol. 13, pp. 131–152, 1984.
- [16] A. Romano and G. Interlandi, "Quale intonazione per il torinese?" in *La fonetica acustica come strumento di analisi della variazione linguistica in Italia*. Il Calamo, 2002, pp. 117–122.
- [17] M. Grice, *The intonation of interrogation in Palermo Italian: implications for intonation theory*. Walter de Gruyter GmbH & Co KG, 2017, vol. 334.
- [18] P. Escudero and D. Williams, "Native dialect influences second-language vowel perception: Peruvian versus iberian spanish learners of dutch," *The Journal of the Acoustical Society of America*, vol. 131, no. 5, pp. EL406–EL412, 2012.
- [19] E. Simon, M. Debaene, and M. Van Herreweghe, "The effect of I1 regional variation on the perception and production of standard I1 and I2 vowels," *Folia Linguistica*, vol. 49, no. 2, pp. 521–553, 2015.
- [20] A.-K. Gartig and G. Rocco, "Soziophonetisches projekt salzburg-triest (sophoprost). untersuchungen zur I1-und I2-phonetik und phonologie junger italienischer deutschlermer," in *Gesprochene (Fremd-) Sprache als Forschungs-und Lehrgegenstand*. EUT-Edizioni Università di Trieste, 2018, pp. 109–137.
- [21] G. P. Georgiou, "Speakers of different I1 dialects with acoustically proximal vowel systems present with similar nonnative speech perception abilities: Data from greek listeners of dutch," *Speech Communication*, vol. 150, pp. 32–40, 2023.
- [22] W. Baker and L. C. Smith, "The impact of I2 dialect on learning french vowels: Native english speakers learning quebeois and european french," *Canadian modern language review*, vol. 66, no. 5, pp. 711–738, 2010.
- [23] S. Lee and M.-H. Cho, "The impact of I2-learning experience and target dialect on predicting english vowel identification using korean vowel categories," *Journal of Phonetics*, vol. 82, p. 100983, 2020.
- [24] P. Boersma, "Praat, a system for doing phonetics by computer," *Glott. Int.*, vol. 5, no. 9, pp. 341–345, 2001.
- [25] A. L. Anwyl-Irvine, J. Massonnié, A. Flitton, N. Kirkham, and J. K. Evershed, "Gorilla in our midst: An online behavioral experiment builder," *Behavior research methods*, vol. 52, pp. 388–407, 2020.
- [26] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2021. [Online]. Available: <https://www.R-project.org/>
- [27] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.
- [28] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest package: Tests in linear mixed effects models," *Journal of Statistical Software*, vol. 82, no. 13, pp. 1–26, 2017.
- [29] R. V. Lenth, *emmeans: Estimated Marginal Means, aka Least-Squares Means*, 2023, r package version 1.8.6. [Online]. Available: <https://CRAN.R-project.org/package=emmeans>
- [30] D. Lüdtke, "ggeffects: Tidy data frames of marginal effects from regression models," *Journal of Open Source Software*, vol. 3, no. 26, p. 772, 2018.
- [31] F. Santiago, P. Mairano, and E. Delais-Roussarie, "Non-native perception of final boundary tones in french interrogatives," in *Speech Prosody 2014-7th International Conference on Speech Prosody*, 2014, pp. 563–567.