Article

LANGUAGE TEACHING RESEARCH

Language Teaching Research
1–27

DOI: 10.1177/13621688241312502
journals.sagepub.com/home/ltr
Sage

Is intonation learnable in the classroom? Evidence from Turkish learners of English

Sinem Sonsaat-Hegelheimer
Iowa State University, USA

John Levis
Iowa State University, USA

Abstract

For language learners, intonation is widely considered to be important in communicating meaning in context, but intonation is also considered by teachers to be difficult to teach, and some have even argued that it may be unteachable. This exploratory study examines whether explicit teaching of three final intonation contours (falling, rising, falling–rising) led to improved perception and production. Thirty-one Turkish learners of English as a foreign language (EFL) participated in a three-week training session on the perception and production of the three contours at the end of a course on English pronunciation. Results from a pre-test/post-test design showed that perception of all three intonation contours improved after instruction, whereas for the production only the falling–rising and rising contours showed improvement. Results also showed that providing contextual information did not affect production but was helpful in perception. This study suggests intonation can improve when it is explicitly taught to L2 learners, like other aspects of pronunciation.

Keywords

EFL, intonation, pronunciation, pronunciation teaching, suprasegmentals

1 Teaching intonation to second language (L2) learners

Since the mid 1900s, intonation has been considered a central part of teaching pronunciation (O'Connor & Arnold, 1961; Pike, 1945), and researchers and teachers have presented dozens of intonation contours as being important for intonation in sentence contexts. Regarding the teaching of intonation, early models such as those put forth for American English by Pike (1945) and for British English by O'Connor and Arnold

Corresponding author:
Sinem Sonsaat-Hegelheimer, Iowa State University, 203 Ross Hall, Ames, IA 50011-2042, USA
Email: sonsaat@iastate.edu

(1961) were simultaneously pedagogical and theoretical. They influenced later pedagogical approaches, such as Allen (1971), who argued that then well-known linguistic descriptions could be used in innovative ways for teaching intonation in sentences and in dialogues. The importance of intonation in language teaching continued during the Communicative Language Teaching era (e.g. Kenworthy, 1987; Wong, 1987), but the number of contours presented as important for teaching was far more limited than the models of the mid-1900s. Brazil et al. (1980), in the development of a theory of discourse intonation, proposed five contours as important for the analysis of natural spoken discourse (rising, falling, rising–falling, falling–rising, level), later developed more fully into the pedagogical model of Discourse Intonation (Brazil, 1997). In contrast, most American pronunciation materials focused only on two final intonation contours – rising and falling – with an emphasis on how the contours were used in particular grammatical structures such as yes/no questions (Levis, 1997).

In a change from sentence-level practice to discourse-level practice, Brazil (1997) argued that intonation finds its fullest meaning in context, and that rising, falling, and level contours have a special role in expressing information structure. This insight thus suggested that the distinctions between different final intonation contours could be most successfully taught in reference to communicative intentions (Thompson, 1995).

Assertions about how to teach intonation through the years assume that teaching intonation leads to learning, but these assertions have not been confirmed by studies demonstrating that this is true. To explore the extent to which intonation is teachable and learnable in a classroom context, this article reports a study of Turkish learners of English as a foreign language (EFL) who were taught to recognize and produce rising, falling and falling–rising intonation. These three intonation patterns were chosen because they are the most common patterns included in teaching materials and because they are relatively common in natural speech (Bolinger, 1986). Rising and falling intonation are central to teaching in all published materials (e.g. Grant, 2016; Lane, 2012), and falling–rising intonation is the next most commonly taught pattern (e.g. Hewings & Goldstein, 1999; Levis & Muller Levis, n.d.). Because it is important for learners to be able to both produce intonation patterns and also to identify and distinguish such patterns in the speech of others, both perception and production were included in teaching and in data collection. However, in this exploratory study, we make no claims about the specific relationship between perception and production.

II Literature review

1 Forms and functions of intonation in English

English intonation involves multiple forms and multiple functions. As to form, intonation features that clearly impact meaning include variations in syllable prominence and pitch movement at the end of a phrase, what we will call final intonation (Levis & Wichmann, 2015). These two aspects of intonation are closely related in that the final prominent syllable in a phrase is also the beginning of final intonation, the focus of our study. Prominence carries a distinct type of meaning from final intonation, as shown in Ladd (1996), where the answer ‘three pounds’ can be accented on either word and

spoken with either rising or falling final intonation (called tune by Ladd) as the appropriate answer to different discourse contexts.

Syllable prominence, which has been widely studied for theory and practice, is the focus of extensive research related to the special role of prominence on the last pitch-accented syllable in a spoken phrase. Prominence, in the Auto-segmental approach to description, is represented by High or Low pitch accents (H*, L*) which may also be imperfectly aligned with the accented syllable (H+L*, H*+L, L*+H, L+H*). Prominence may function to signal the information structure of speech (e.g. Baumann & Grice, 2006) or semantic contrasts (Cummins & Rohde, 2015), among other functions.

Finally, intonation at the end of a phrase is often referred to in teaching materials as intonation, as final intonation or as an intonation contour. The most readily identifiable forms of final intonation in English (not including the nuclear stressed pitch accent) include rising, which ends with H% boundary tone or falling, which ends with L% boundary tone before a phrase break (Nolan, 2022), as well as combinations of these movements, e.g. falling–rising (LH%).

In regard to function (i.e. meaning), falling pitch may signal that an assertion is completed or closed, while rising pitch is often heard as communicating openness, that is, that the utterance is incomplete as part of discourse and requires a response (Dalton & Seidlhofer, 1994; Halliday, 1967; Levis & Wichmann, 2015). In discourse, intonation has been associated with the negotiation of meaning and contributions to information structure (Pickering, 2018) and, indeed, intonation's strongest contribution to communication may be in the construction and co-construction of spoken discourse (Chun, 1988).

Meaning differences associated with intonation are also evident at the sentence level. For example, the sentence 'They're ready to go' spoken with falling intonation may simply assert information to the interlocutor, but spoken with rising intonation, it will be heard as questioning the assertion. The same grammatical structure with a falling–rising intonation will be heard as having a 'but' about it as in 'They're ready to go (but) . . .' (Halliday, 1967). If the speaker does not continue to speak beyond the end of the utterance, the falling–rising intonation implies that something is left unsaid. This implicature might involve uncertainty or incredulity (Ward & Hirschberg, 1986). However, if the falling–rising intonation is immediately followed by more speech, it will simply be interpreted as a signal that the speaker is not yet finished.

2 Turkish intonation

The intonation of Turkish is less well described than that of English, especially in regard to utterance-level prosody. At the word level, Turkish is considered to be a pitch accent language, in which every stress at the word level usually occurs at the end of the word and is signaled by H* (high pitch accent) or !H*, a downstepped high pitch accent in a series of stressed syllables (Ipek & Jun, 2013). Turkish is also an agglutinative language in its morphosyntax, with multiple possible suffixes attached to the root. The pitch accent falls at the end of morphosyntactically complex words, except in the case of certain suffixes that cannot be accented. In such cases, the accent is on the root of the word. Turkish is also syntactically verb final, but final verbs are

not accented (Ipek, 2015) and spoken at a low pitch. In shape, then, most utterances may sound like they have falling intonation.

However, researchers have argued that utterance-level prosody is constrained by the morphosyntactic complexity of Turkish. Kabak (2016) says that 'Turkish has . . . a very poor inventory of utterance level pitch events' and 'does not employ prosodic correlates of information structural units in the way that languages such as English and German do' (p. 17). Other researchers (Ipek & Jun, 2013) have argued that rising intonation signals the intermediate phrase boundary after the L pitch of the nuclear word (the final verb), creating a LH (rising or falling–rising) boundary tone. The LH boundary tone is not attested for final boundary tones, which Ipek (2015) argues are either L% (falling, continuing the pitch of the final verb) or H% (rising). H% boundary tone, though described by Ipek (2015), is less well attested. As a result, Turkish appears to favor final falling intonation because of the compression of the pitch accent on the final verb. Acoustically, falling–rising pitch shows up in non-final phrases (H*, L final verb, H) but is not attested at final boundaries. The presence of rising intonation (with H% boundary tone) is less well described. From this research, we would expect Turkish learners to be skilled at identifying and producing falling intonation, but there is less confidence about how well they will perceive and produce falling–rising and rising intonations.

3 Teaching intonation

Despite the long history of intonation being considered an essential aspect in teaching pronunciation, teachers regularly report that intonation is particularly difficult to teach (Burri et al., 2017; Couper, 2016; Sonsaat, 2017). Couper (2019), in a study of New Zealand teachers of English as a second language (ESL), reported that they felt less confident about teaching stress and intonation than they did segmentals. Kochem (2022), in a study of online teacher training for pronunciation, reported that teachers wanted knowledge of how to consider 'the context and the speaker's intent' when working with intonation because intonation seemed 'very subjective rather than tied to a certain set of rules governing how to do it' (p. 69). More troubling is the assertion of Jenkins (2000), who said that because of its complexity in communicating meaning, discourse intonation is largely 'not teachable' (p. 152) to speakers of English as a lingua franca. By not teachable, Jenkins explained that the use of classroom time did not lead to acquisition because accurate rules about intonation are so difficult to provide. Pennington and Ellis (2000) provided some evidence for this assertion. In their study of Cantonese speaking learners of English, the researchers found that learners did not pay attention to how pitch communicated meaning differences, and that even when they were instructed in how English uses pitch to communicate meaning, the learners still did not perceive meaning for intonational features.

There is limited research showing positive impact of training on perception and/or production of intonation. There is also research that showed no impact of training (Tanner & Landon, 2009) or confusion in learners' perception of some intonation contours (Hori et al., 2021). Furthermore, studies such as Pennington and Ellis (2000), which suggest that learning does not follow teaching, may only indicate that relatively brief explanations of how pitch communicates meaning do not provide any evidence for whether final

intonation is learnable with adequate teaching and practice in a classroom context. We therefore need more experimental studies of intonation instruction to determine if instruction of intonation leads to improvement.

Research has shown that teaching suprasegmentals in general leads to improved comprehensibility for second language (L2) speakers (Gordon & Darcy, 2016; Zhang & Yuan, 2020). Improved comprehensibility means that listeners find it easier to listen to, and understand, L2 speakers. Recent evidence also indicates that intonation improvement has a positive influence on L2 speakers' perceptions of their own intelligibility (LaScotte et al., 2022). In a final finding that demonstrates the importance of intonation for comprehensibility improvement, Jiang and Chun (2023) looked at how the spontaneous speech of Mandarin learners of English in China and the US improved after four weeks of discourse intonation training using visual feedback. Training improved both groups' comprehensibility, showing that intonation is teachable, and intonation training can result in some aspects of global improvement for spontaneous speech.

In studies that show how formal instruction can improve the intonation of L2 learners, two papers involving English and French are instructive. Hardison (2004) found that employing extensive perception and production training led to learners becoming more automatic in producing and perceiving intonation contours in L2 French. Herment and Tortel (2021), arguing that French learners of English use too many final rising intonation contours (a transfer from French) instead of falling contours to indicate finality and falling–rising contours to indicate non-finality, highlight the importance of intonation in discourse, in understanding whether a speaker is done with their turn in speech or not. Thus, the authors developed a visualization technique to help L2 learners understand the difference between English and French intonation systems.

Pedagogical treatments of intonation typically focus on a few contours, especially the falling, rising, and falling–rising contours. Analysis of L2 speech has shown that rising and falling intonation both occur regularly on all grammatical structures but express different meanings (Fries, 1964; Geluykens, 1988; Hedberg et al., 2017; Shepherd, 2011; Thompson, 1995; Warren, 2016). Unfortunately for teachers, there are few studies examining whether intonation improves as a result of instruction (de Bot & Mailfert, 1982; Goh, 1994; Hori et al., 2020; Le & Brook, 2011). It is these few studies that are central to our study. De Bot and Mailfert (1982) found that intonation perception training led to production improvement for L2 learners of English. Goh (1994) took a discourse intonation approach to teaching intonation to 16–17 year-old students in Singapore. Students received three weeks of formal instruction that focused on perception and production of final intonation. Results showed that the learners heavily used level and falling intonation but had difficulty perceiving and using falling–rising intonation. Le and Brook (2011), in a small-scale study of English learners, explored the use of Praat (a speech analysis software; Boersma & Weenink, 2024) as a tool to visualize intonation. They reported that learners were able to produce question forms more clearly and naturally after instruction, a standard that was based on the researchers' impressionistic analyses of pitch tracings on Praat. In a recent study of Japanese learners of English that is most similar to our study, Hori et al. (2020) examined the intonation improvement of Japanese learners of English and found that perception was easiest for falling contours, followed by rising contours, and then by falling–rising contours.

Previous studies have also argued that context plays a crucial role when teaching and learning intonation. The importance of context is a core assumption of a discourse approach to teaching intonation (e.g. Chun, 1988; Pickering, 2018), of pedagogical critiques of intonation teaching materials (Levis, 1999), and of the importance of awareness of intonational meaning (Liu et al., 2024). In a study of how contextual aspects of prominence promoted L1-related transfer, Ortega-Llebaria and Colantoni (2014) found that context helped Spanish first language (L1) learners of English to express meaning when retelling a story for children, as opposed to a comparison group who repeated sentence intonation. Levis and Pickering (2004) demonstrated that English speakers reading the same sentences with and without context used different intonational cues. However, no studies have examined whether the impact of classroom instruction on the perception and production of intonations varies when context is provided.

In this exploratory study, we investigate the impact of classroom instruction on L2 speakers' perception and production of three intonation contours: falling (F), rising (R), and falling–rising (FR). There is very little research on how language learners improve their perception and production due to classroom instruction.

This study addresses two research questions. Research question 1 addresses whether intonation is like other aspects of pronunciation (i.e. if it is taught, will it be learned?), and whether the effects of teaching apply both to perception and production. Research question 2 addresses the effects of contextual information on learning. Because intonational meaning is sensitive to discourse context, we look at the effects of limited contextual information on the success of intonation learning.

- Research question 1: What is the impact of explicit instruction on perception and production of intonation contours?
- Research question 2: What is the impact of context on the perception and production of intonation contours (F, R, FR)?

Our predictions for this study were that the perception of intonation contours would show a consistent improvement whereas production would not show the same levels of improvement. We also expected that the falling–rising intonation would show the largest amount of improvement among learners for two reasons: First, we expected it to be the most unfamiliar contour and so to start at a lower level of accuracy, and second, because it is a more complex contour that includes pitch movement in two directions. Finally, we expected that greater provisions of contextual information would result in larger improvements for both perception and production. Intonation is a discourse level pronunciation feature, and it makes sense that even minimal discourse-level information could assist learners in connecting intonational form and meaning.

Before describing the study in detail, it is important to state that we did not follow a theoretical framework in teaching intonation to these learners. The most consistent theoretical framework for pedagogical approaches to intonation is Discourse Intonation, or DI (Brazil, 1997; Pickering, 2018). DI provides a theory of intonational meaning (i.e. information structure) that is central to connecting form and function, but this theory of meaning was not relevant to our goals in this study. In addition, DI's descriptive apparatus for intonation (including tone unit markings, prominent syllables, and directional arrows

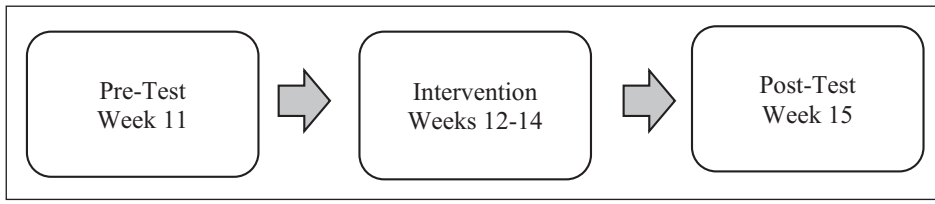

Figure 1. Design of the study.

preceding but not following the utterance) was more complex than was desirable in our study. In addition, there is evidence that DI's markings are not easy for learners to interpret (Niebuhr et al., 2017). Instead, we chose to use a pedagogical approach that is common in both British (e.g. Hewings, 2018) and American materials (e.g. Grant, 2016) by using directional arrows to represent rising, falling, and falling–rising intonation. Additionally, when learners practiced with longer text to mark intonation, they worked on breaking the text into thought groups first, which they had learned previously, and then worked on marking intonation.

III Methods

This study follows a quasi-experimental design with a pre-test, training intervention, and a post-test (Figure 1). Because of the exploratory nature of the study and because the instructor left the school after the study period, we did not conduct a delayed post-test. A control group was not used because we did not want to deny a potentially beneficial treatment on this aspect of pronunciation to control group participants. Furthermore, an additional group was not available. Finally, our goal was to demonstrate the relative development of different patterns under different contextual conditions.

The data presented in the study is all quantitative. Participants in the study were 31 pre-service English teachers whose L1 is Turkish (M_{age} = 20, range = 18–21 years). All participants in this study (N = 35) were enrolled in a 16-week pronunciation class (2.5 hours per week) at a large private university in Turkey. In this class, students learned segmentals (i.e. vowels and consonants) between weeks 1–6 and suprasegmentals (i.e. word stress, rhythm, thought groups, prominence, and intonation). Between weeks 7–14, instruction focused on suprasegmental features, and in weeks 12–14 (three weeks), instruction focused only on teaching of intonation. Pronunciation instruction was provided by one of the researchers in this study. This was the first pronunciation class for the students; they did not receive any training on intonation or any other prosodic feature before this class. Students took a pre-test measuring their perception and production of intonation in Week 11 and took the post-test in Week 15 after the teaching of intonation. After student grades were submitted, the instructor asked students if they would give consent for their work to be included in research. All 35 students agreed for their class data to be used for this research, but we only used the data from 31 students since the other four students missed at least one class session between weeks 12–14. The study was approved by the Human Research Ethics Committee of the university where the data were collected.

1 Procedures

As part of the pronunciation class, students took a diagnostic test at the beginning of the semester and a final assessment test at the end. Information from this assessment is not included in this study. Throughout the semester, students had various recording tasks for practice and assessment. Depending on the topic and class procedures, some of these were recorded and uploaded to the course platform by students. For instance, those that were used for peer feedback and self-assessment purposes were self-recorded by students, but others were recorded by the instructor. For intonation, students met the instructor in week 11 to record production test items and took the perception test during class time in the same week (pre-test). The same procedures were completed in week 15 for the post-test. The recordings and perception tests were later rated by both the instructor (first author) and the second author in this study. To mitigate potential bias we might have had as the researchers, we listened to pre- and post-test productions in random groupings. The number of recordings made it impossible to make comparisons or think about which test's recordings we were listening to.

a Training intervention. The three weeks of explicit intonation instruction focused on teaching the forms and functions of three intonation contours (i.e. falling, rising, falling–rising). The instructor modeled and taught intonation using General American English as spoken in the US Midwest. Although the contours and meanings we taught would have been the same in British English, we did not model this variety because neither researcher has a command of spoken British English, nor is British English an enforced model for the Turkish learners. In this study, we did not follow a traditional, grammar-based teaching approach. Rather, we showed the connection between intonation contours (form) and the meanings (function) they could convey in a given context. We emphasized that a speaker could use any of the three intonation contours with the same set of words to produce different communicative meanings. Both perception and production activities in class progressively involved the identification and production of intonation contours in single words, short phrases and conversations. Students had to pay attention to meaning, that is, to the function of the intonation choices.

In weeks 12 and 13, students were provided with explicit guidelines for the forms and functions of the three intonation contours. Practice moved from tightly controlled to freer, more communicative practice. Different intonation choices were considered when completing the activities, and the idea that the intonation choices depend on the meanings to convey was highlighted. This was the case for the perception activities in which students would need to identify the intended meaning of the speaker among multiple options based on the intonation contour. In larger texts where students had more contexts, they also identified intonation contours at the end of thoughts groups to be able to tell if a statement was complete or incomplete. Students also had communicative tasks in which they responded to given situations with a short utterance by paying attention to the intonation contours at the end of their thought groups. Additionally, in week 13 students practiced identifying and imitating the intonation patterns they heard in a TED talk entitled 'What causes insomnia?' This TED Talk was chosen because its intonation contours exemplified some of the functions of

intonation (i.e. listing a series of items, finality) taught in class and it would give students the opportunity to work with authentic speech.

Week 14 included review of the contents covered in Weeks 12 and 13, followed by one-minute speeches by class members. For this speech, students were given 5 minutes of preparation time during which they chose their topic and prepared what they wanted to say. The instructor highlighted key aspects that had been developed during the intervention, such as the pitch movement at the end of thought groups and the communicative impact that those choices may have on meaning. Students were not given a particular model to imitate for the free speech task, but they were told to think about the intonation of the TED Talk and other previous activities to remember how the meaning of what was said might change depending on the final intonation. While each student was delivering their speech, other students were asked to monitor their classmates' use of intonation and provide feedback on whether their use of intonation was appropriate for the meaning they were trying to convey. For example, students often talked about the importance of the falling intonation at the end of finished statements. Appendix A shows the weekly content and types of exercises.

b Pre-test and post-test. The data in the study come from the tasks in students' perception and production tests at pre-test (Week 11) and post-test (Week 15). There were three parts in both perception and production tests with 43 items in total (Appendix B). Perception and production test items were identical. Students always completed the production tasks first so that their performance would not be affected by what they heard in the perception tasks.

In the production test, students were given instructions at the beginning of each part. In Part 1, they were told to read each of the 15 single-word statements with a given intonation contour, indicated by an arrow. In Part 2, students were given 16 multiple-word phrases along with the intended meaning they had to convey by producing the phrases. Students were expected to use the intonation contour that would be appropriate for the intended meanings. In Part 3, students were given 12 short dialogues including the utterances for Speakers A and B. They were asked to say Speaker B's utterance by looking at the intended meaning given on the slide and the context provided by Speaker A's utterance (for examples, see Figure 2).

In Part 1 of the perception pre-test, students listened to 15 single-word statements with two exceptions (i.e. 'sort of' and 'like it') and chose the intonation contour they thought they heard: falling, rising, or falling–rising. In Part 2, they listened to 16 multiple-word phrases and chose the answer that best described the intended meaning. Each phrase had two potential meanings, one showing certainty (with falling intonation) and one showing reservation (with falling–rising intonation). For example, when listening to the utterance 'I'd love to', students could choose either (a) '. . . and I will' or (b) '. . . but I cannot'. In Part 3, students listened to 12 short dialogues and were asked to choose the answer that best described the intention of the second speaker based on the intonation contour used by the speaker. For example, hearing the utterance of Speaker B in the following dialogue example, students could choose either (a) I'd be glad to help you or (b) I'd love to, but I really don't have time. Option A was produced with falling intonation whereas option B was produced with falling–rising intonation.

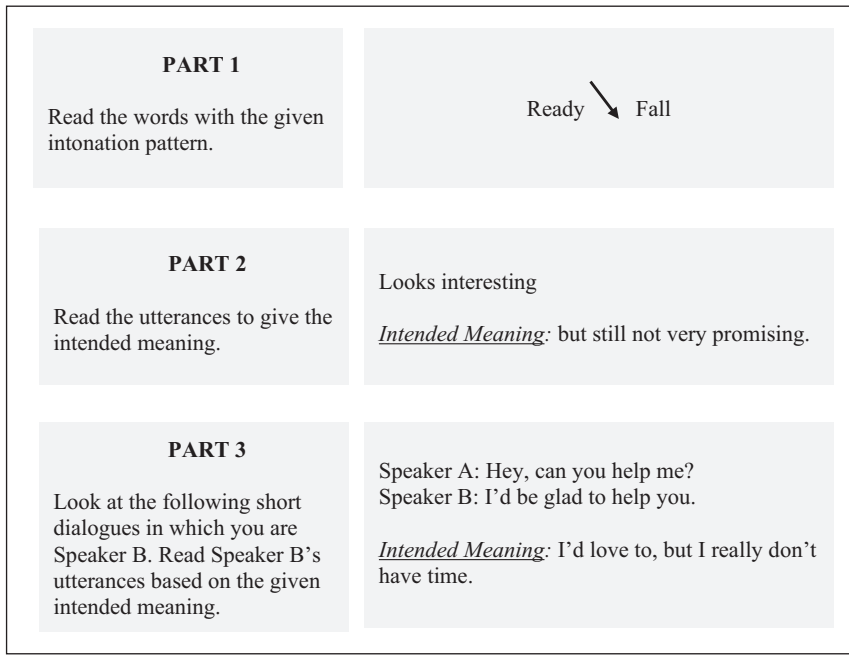

Figure 2. Presentation of production test items (One example from each part).

Example dialogue from Part 3:

Speaker A: Hey, can you help me?
Speaker B: I'd be glad to help you.

Students took the perception test on Moodle Learning Management System (LMS) both times. Items in the perception test were produced by a male and a female, both North American English speakers. The items in each part of the pre-test were presented in a random order since each part had repeated items with different meanings based on their intonation contours.

2 Data analysis

The data used to answer the research questions are the scores from the pre- and post-test perception and production tests. In the perception test, we coded incorrect identification of the intonation contours as 0 and correct identifications as 1. For the production test items, we first identified the intonation contours used by the students for each item at both pre- and post-test in Part 1 ($N=15$), Part 2 ($N=16$), and Part 3 ($N=12$). In total, we listened to 2,666 tokens for 43 items produced by 31 students at pre- and post-test (i.e. $(43 \times 31 \times 2)$. We chose to use a listener-based analysis because we were evaluating categorical variables, which required evaluation by listeners who could determine whether productions matched the phonological categories or rising, falling and falling–rising.

Such listener-based evaluations are essential to pedagogical pronunciation research that evaluates phonological categories (Derwing & Munro, 2015). An acoustic analysis could tell us only about pitch range variations but not whether the productions were heard as matching particular phonological categories.

The Cohen's kappa coefficient, calculated by using the *irr* package (Gamer et al., 2022) in R for the rating of 1,333 items in the pre-test ($k = 0.79$) and post-test ($k = 0.72$), showed substantial agreement between the two coders. We discussed the items we did not agree on to come to an agreement. On several occasions where the intonation contour used by a student was uncertain, we consulted Praat (Boersma & Weenink, 2024). Once we had agreed on the intonation contours students produced (i.e. falling, rising, falling–rising), we replaced them with 0 or 1 based on whether students the correct contour according to the prompts and intended meanings.

The statistical analyses were performed using R software (version 4.2.3; R Core Team, 2023). We ran binomial distribution mixed effects models with a logit link for both perception and production items. The models were fitted using the *lme4* package in R (Bates et al., 2015). We had two models for both perception and production data. In the first models for each data set, we analysed the main effect of training. In these models, Time (i.e. pre-test and post-test) was the fixed effect, and we had participants and items as random effects since the study has a repeated measures design. To assess the statistical significance of the fixed effects in the models, we relied on the *z*-values associated with each coefficient estimate. We used the *emmeans* package in R (Lenth, 2023) to calculate the estimated marginal means. The data wrangling and visualization were performed using *tidyverse* (Wickham & Grolemund, 2017) and *janitor* (Firke, 2023) packages in R. To check the model assumptions, we used the *DHARMa* package (Hartig, 2022) and to convert the log odds to odds ratios, we used the *sjPlot* package (Lüdecke, 2023).

In the second models for perception and production, we included interaction between Time (i.e. pre-test, post-test) and Contour (i.e. F: Falling, FR: Falling–Rising, and R: Rising). Additionally, we included Context (Context 1: single word statements, Context 2: multiple-word statements with an intended meaning, and Context 3: a short dialogue with an intended meaning) as a controlling predictor. In both first and second models, we did dummy coding by hand (a type of contrast coding) for the fixed effect, Time (0 = Time 1, 1 = Time 2) (Winter, 2020).

We included the interaction between Time and Contour because we hypothesized that the effect of training might be different on each intonation contour in this exploratory study. Context is a controlling predictor variable in this study, not the main effect, because research in intonation shows that contextual information could be influential in students' perception and production of intonation contours (Wichmann, 2015). We did not expect context to have an interaction with time because if it had an impact, the effect of context would be applicable to both times.

IV Results

1 Perception of intonation

We calculated the descriptive statistics by Time and Contour and Time and Context. Table 1 shows that all intonation contours improved from pre-test to post-test. Standard

Table 1. Descriptive statistics for perception accuracy by Time and Contour.

Contour	Time 1		Time 2		Overall	
	M	SD	M	SD	M	SD
Falling	0.71	0.45	0.87	0.32	0.79	0.40
Falling–Rising	0.70	0.45	0.80	0.39	0.75	0.42
Rising	0.58	0.49	0.76	0.42	0.67	0.47

Table 2. Descriptive statistics for perception accuracy by Time and Context.

Context	Time 1		Time 2		Overall	
	M	SD	M	SD	M	SD
Context 1	0.67	0.47	0.81	0.38	0.74	0.43
Context 2	0.62	0.48	0.80	0.40	0.71	0.45
Context 3	0.81	0.38	0.89	0.30	0.85	0.35

deviations (shown between parentheses in Table 1) were smaller in the post-test. Despite the fact that variability remained substantial, the results suggest less individual variability after the training. Table 1 also shows that the rising intonation was the most challenging for the students at the beginning of the study and showed the largest improvement from the pre-test to the post-test. However, at the post-test, rising intonation was still identified least successfully.

Table 2 shows students best identified intonation contours correctly in Context 3, in which the most extensive context was provided for them. However, the gap between Context 3 and other contexts is smaller at the post-test, which may suggest that students relied more on context for the intonation contours they were not familiar with at the beginning of the study. Another important point is that the standard deviations became smaller at the post-test, which suggests that training resulted in a decrease of variability.

We fit two logistic mixed-effects models to the perception data, one to estimate the overall effect of training from pre-test to post-test (Time), averaging across all intonation contours in all three contexts we have in the study; and a more complex model to show the interaction between the Time and Contour (i.e. falling, falling–rising, and rising) and the impact of Context.

According to the first model's results shown in Table 3, the odds ratio for the intercept was bigger than 1.00, suggesting that it was more likely than not that students identified the intonation contours correctly at the pre-test. The odds ratio for Time >1.00 and its 95% CI indicate that students' likelihood of identifying the intonation contours at the post-test was higher than at the pre-test, showing improvement in perception accuracy after the training. The model accounted for 6% of the variation in the perceptual accuracy of intonation (marginal $R^2=0.060$) based on the fixed effects alone whereas fixed and random effects together explained 40% (conditional $R^2=0.396$) of the variation. Looking at the *DHARMa*-simulated residuals, we saw that there were no problems with the model fit.

Table 3. Summary of logistics mixed-effects model fit to perception data.

Fixed effects	OR	*SE*	*z*	95% CI	*p*
Intercept	2.88	0.60	5.11	[1.92–4.33]	< .001
Time	3.14	0.67	5.40	[2.07–4.76]	< .001
Random effects		*Correlation*			
Participants:					
Intercepts	0.67				
Slopes: Time	0.96	−0.13			
Items:					
Intercepts	0.99				

Notes. Model syntax: glmer (correct ~ time + (1 + time | participant) + (1 | item), family = binomial). Odds ratios were computed using the tab_model function of the *sjPlot* package (Lüdecke, 2023). OR = odds ratio; SE = standard error; CI = confidence interval.

In our second model, we included Time × Contour interaction as well as Context as a controlling predictor. We included the interaction between the Time and Contour because we did not expect each intonation contour to show the same level of improvement. To the best of our knowledge, there is no research looking at the development of perception and production of different intonation contours for Turkish-speaking or other L2 learners of English; however, our descriptive statistics in both perception and production of intonation contours encouraged us to keep the interaction of Time and Contour in our analysis. Because Context is a controlling predictor variable in the model, we initially fit the model also to include the interaction of Context with Time; however, including this interaction did not improve the model fit ($\chi^2 = 0.3091$, df = 2, $p = .85$). Therefore, we went with the model which included only the interaction between Time and Contour while keeping Context as a controlling predictor variable. We included by-participant and by-item random intercepts and by-participant random slope for Time. Initially, we fit a model not only with by-participant Time random slopes but also for the Contour and Context; but these models did not converge even when we followed the methods (i.e. centering the predictor, increasing the iteration number, and trying alternative optimizers) recommended by Brauer and Curtin (2018). Therefore, we included only by-participant random slopes of Time, which is also more important in the design of this study. Based on the examination of the model assumptions using *DHARMa*-simulated residuals, there were no problems with the model fit.

According to the results of this model, there was a significant interaction between Time and Contour, likely due to the results for the falling–rising contour (OR = 0.74, SE = 0.08, CI = 0.59, 0.92, $p = .007$). Based on this model, we ran pairwise comparisons with the Tukey method using the *emmeans* package in R (Lenth, 2023). Results of the pairwise comparisons are presented in odds ratio (OR). In pairwise comparisons, if odds ratio in comparisons is equal to 1, there is no difference between the two compared values; if the odds ratio is bigger than 1, the odds of the first value (Time 1 on the left) is higher than the odds of the second value (Time 2 on the right); and if the odds ratio is between 0 and 1, the odds of the second value is higher than the odds of the first value.

Table 4. Pairwise comparison of estimated accuracy in perception between pre-test and post-test.

Contours	OR	SE	95% CI	*p*-value
Time 1 F – Time 2 F	0.22	0.05	[0.11, 0.47]	< .0001
Time 1 FR – Time 2 FR	0.42	0.10	[0.21, 0.83]	.004
Time 1 R – Time 2 R	0.30	0.10	[0.11, 0.79]	.005

Notes. Based on model syntax: glmer (correct ~ scale(time)*contour + context + (1 + scale(time) | participant) + (1 | item), family = binomial, data = intonation.perception); F = falling; FR = falling–rising; R = rising; OR = odds ratio; SE = Standard error; CI = Confidence interval.

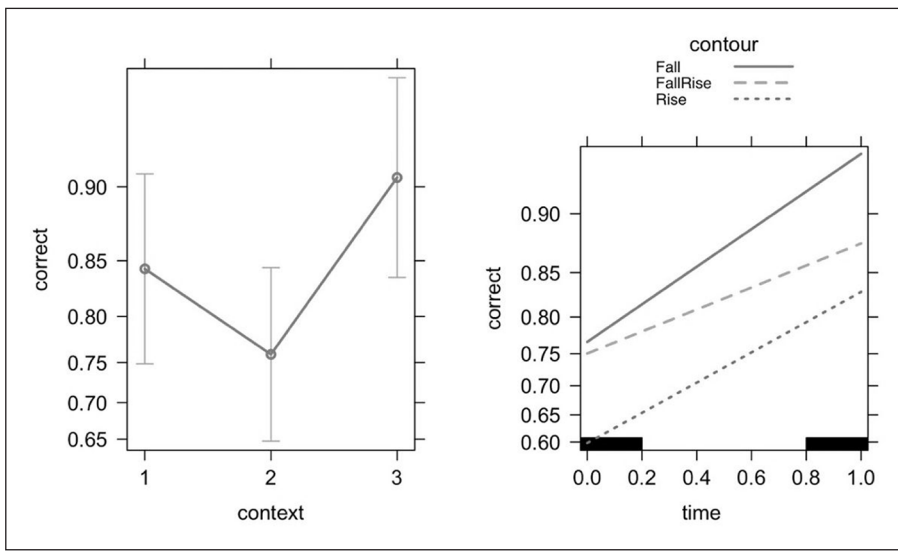

Figure 3. The Impact of Context Averaged Across Contours and Time (Left) and Interaction between Contours and Time (Right) – Perception.

As shown in Table 4, the ORs for all comparisons are between 0 and 1, CIs not including 1.00, meaning that the odds of identifying all intonation contours correctly at the post-test was significantly higher than identifying them correctly at the pre-test. These results suggest that the training was influential for the perception of all intonation contours (Figure 3, right).

As for the Context effect, pairwise comparisons of context averaged across the Time and Contour show no significant difference between Context 1 and Context 2 (OR = 1.70, SE = 0.64, CI = 0.71, 4.10, $p = 0.32$) and between Context 1 and Context 3 (OR = 0.56, SE = 0.22, CI = 0.21, 1.45, $p = 0.33$). However, there is a significant difference between Context 2 and Context 3 (OR = 0.33, SE = 0.11, CI = 0.14, 0.76, $p = 0.006$), suggesting that the odds of identifying intonation contours correctly was higher in Context 3 compared to Context 2 (Figure 3, left).

Table 5. Descriptive statistics for production accuracy by Time and Contour.

Contour	Time 1		Time 2		Overall	
	M	SD	M	SD	M	SD
Falling	0.86	0.34	0.87	0.32	0.87	0.33
Falling–Rising	0.14	0.35	0.44	0.49	0.29	0.45
Rising	0.34	0.47	0.66	0.47	0.50	0.50

Table 6. Descriptive statistics for production accuracy by Time and Context.

Contour	Time 1		Time 2		Overall	
	M	SD	M	SD	M	SD
Falling	0.42	0.49	0.63	0.48	0.52	0.49
Falling–Rising	0.52	0.50	0.66	0.47	0.59	0.49
Rising	0.52	0.50	0.68	0.46	0.60	0.48

2 *Production of intonation*

For production data, we first calculated the descriptive statistics by Time and Contour and Time and Context. Table 5 shows a ceiling effect for the falling intonation, which was generally produced correctly at the pre-test and showed no improvement at the post-test. Falling–rising and rising intonation patterns showed substantial improvement from pre- to post-test. As for Context (Table 6), Context 1 seemed to be the least helpful at the pre-test but there is not much difference between the three contexts at the post-test.

As in the perception analysis, we fit two mixed effects models to analyse the production data. The first logistic mixed effects model includes Time as the main fixed effect with by-participant intercepts and slopes and by-item intercepts. The odds ratio for the intercept < 1.00 (Table 7) shows that it was less likely for students to produce intonation contours correctly at the beginning of the study. However, the 95% CI includes 1.00; therefore, this finding was not significant. The odds ratio for Time > 1.00 and its 95% CI, which does not include 1.00, indicate that after the training, students were more likely than not to produce intonation accurately (that is, the improvement significant). The model accounted for 5% of the variation in the production accuracy of intonation (marginal $R^2 = 0.052$) based on the fixed effects alone; whereas fixed and random effects together accounted for 53% (conditional $R^2 = 0.531$) of the variation. Upon examining the model assumptions using *DHARMa*-simulated residuals, all tests showed that the model effectively represented the dataset.

Our second model for the production data included Time x Contour interaction and Context as a controlling predictor. We examined the interaction between Time and Context for production analysis as well, but this did not improve the model fit ($\chi^2 = 1.6532$, df = 2, $p = 0.43$). We therefore selected the model which included only the interaction between Time and Contour while keeping Context as a controlling predictor variable. This model

Table 7. Summary of logistics mixed-effects model fit to production data.

Fixed effects	OR	*SE*	*z*	95% CI	*p*
Intercept	0.92	0.25	−0.29	[0.54–1.57]	.770
Time	3.34	0.62	6.50	[2.32–4.81]	< .001
Random effects		*Correlation*			
Participants:					
Intercepts	0.05				
Slopes: Time	0.85	1.00			
Items:					
Intercepts	1.71				

Notes. Model syntax: glmer (correct ~ time + (1 + time | participant) + (1 | item), family = binomial). Odds ratios were computed using the tab_model function of the *sjPlot* package (Lüdecke, 2023). OR = odds ratio; SE = Standard error; CI = Confidence interval.

Table 8. Pairwise comparison of estimated accuracy in production between pre-test and post-test.

Contours	OR	SE	95% CI	*p*-value
Time 1 F – Time 2 F	0.79	0.16	[0.38, 1.38]	.72
Time 1 FR – Time 2 FR	0.19	0.03	[0.10, 0.34]	< .0001
Time 1 R – Time 2 R	0.22	0.06	[0.10, 0.51]	< .0001

Notes. Based on model syntax: glmer (correct ~ scale(time)*contour + context + (1 + scale(time) | participant) + (1 | item), family = binomial, data = intonation.production). F = falling; FR = falling–rising; R = rising; OR = odds ratio; SE = standard error; CI = confidence interval.

included by-participant and by-item random intercepts and by-participant random slope for Time. The model with more complex random effects structure, by-participant random slopes of Time, Contour, and Context, did not converge even when we followed the methods (i.e. centering the predictor, increasing the iteration number, and trying alternative optimizers) recommended by Brauer and Curtin (2018). Therefore, we went with the model in which we included only by-participant random slopes of Time. After inspecting the model assumptions using DHARMa-simulated residuals, all examinations indicated that the model accurately captured the characteristics of the dataset.

According to this model, there was a significant interaction between Time and the falling–rising contour OR = 1.96, SE = 0.24, CI = 1.55–2.48, $p = 0.001$ and rising contour (OR = 1.79, SE = 0.28, CI = 1.32–2.43, $p = 0.001$). We ran pairwise comparisons using the Tukey method using the *emmeans* package in R (Lenth, 2023). Results of the pairwise comparisons are presented in odds ratio (OR). According to the results in Table 8, even though Falling intonation had more likelihood of being produced correctly at the post-test because of < 1 pairwise comparison of Time 1 and Time 2 F, the 95% CI including 1 does not indicate significant improvement. This was expected based on descriptive findings. However, both falling–rising and rising intonation contours were significantly more likely to be produced correctly at post-test, suggesting positive impacts of training on the production of these contours (Figure 4, right).

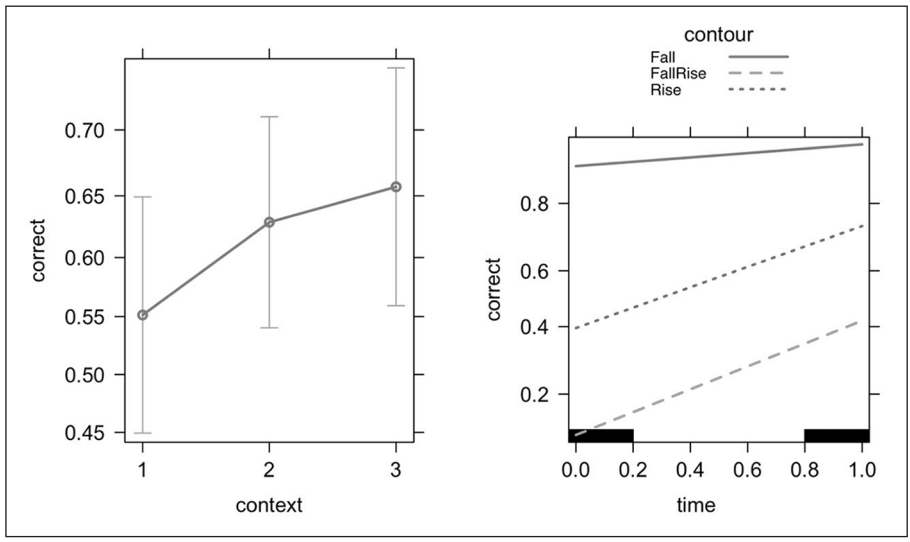

Figure 4. The Impact of Context Averaged Across Contours and Time (Left) and Interaction between Contours and Time (Right) – Production.

The ORs obtained from the pairwise comparisons of Context averaged across the Time and Contour suggested the favorable order of Contexts as 1, 2, and 3 respectively; however there was no significant difference between Context 1 and 2 (OR = 0.72, SE = 0.19, CI = [0.38, 1.37], $p = 0.45$), Context 1 and 3 (OR = 0.64, SE = 0.18, CI = [0.32, 1.26], $p = 0.26$), and Context 2 and 3 (OR = 0.88, SE = 0.22, CI = [0.49, 1.59], $p = 0.87$). These results suggest that no context was particularly helpful for students to produce intonation more correctly (Figure 4).

V Discussion

Our primary goal was to determine if explicit teaching of final intonation led to improvement, as is the case for many other features of pronunciation (Lee et al., 2015). We included three intonation contours in our instructional intervention, falling, rising and falling–rising. Our study posed two research questions, each addressing the development of the perception and production of intonation though explicit instruction.

Research question 1 addressed the degree to which explicit instruction improved the perception and production of the three contours. The results for perception showed significant improvement for all contours from pre-test to post-test. Descriptive statistics showed modest improvement for the falling and falling–rising contours and a more substantial improvement for the rising contour, which the participants/students handled with the lowest accuracy. The results for production gave different patterns for different contours: the falling contour, which displayed the highest levels of accuracy in the pre-test, stayed the same at post-test. The rising and falling–rising contours both showed large and significant improvement at post-test.

A study that resembles ours, Hori et al. (2021), showed improvement for perception of the falling–rising contour but not the other contours. In contrast, we saw improvement for all three contours in our study. Hori et al. and our study were carried out in an EFL environment (Japan, Turkey), and both involved college-age students who were English majors. Hori et al. found that falling–rising intonation was the most difficult to perceive initially but that it also showed the greatest levels of improvement. In contrast, both the falling and rising intonation showed no improvement at post-test. Similarly, the falling–rising contour was the most challenging for the participants. Nonetheless, the results reveal improvement at the level of perception for all three contours. This may indicate that while perceiving intonation may be challenging for L2 learners, it can improve with explicit instruction.

Research question 2 explored to what extent providing meaningful context influences the perception and production of intonation. A basic assumption of discourse approaches to intonation is that the communicative value of intonation becomes most evident when there is a context that helps understand why a speaker has made a particular intonational choice; there are convincing arguments that intonation learning may show greater intake when such learning is connected to awareness of meaning in context (Brazil, 1997; Chun, 1988; Pickering, 2018). In this study, we explored the influence of context in both the development of perceptive and productive skills. Our results indicate that access to contextual information, however minimal, was helpful for the perception of final intonation, confirming Thompson's (1995) assertion that communicative intentions are important in distinguishing rising and falling intonation. Contextual information seems to have facilitated the perception and production of intonation. This suggests that L2 learners may be more sensitive to meaning differences once they understand how meaning differences relate to how something was said, but that production of intonation lags behind perception. This is also in line with recent research showing that awareness of how intonation affects meaning may be essential to learning to use intonation in discourse (Liu et al., 2024). In addition, we deliberately limited our manipulation of context so that we could explore how much contextual information was needed to improve learners' final intonation. Other studies have demonstrated that a focus on contextual meaning can lead to changes in perception and production (Ortega-Llebaria & Colantoni, 2014), and that there are differences in the pragmatic force of contour choice (Thompson, 1995).

Our results reveal that, like other phonological features, final intonation can be learned. These findings are confirmed by other studies. For example, Hardison (2004) demonstrated that English-speaking learners improved their perception and production of French intonation using technology-assisted instruction over three weeks of training. De Bot and Mailfert (1982) and Jiang and Chun (2023) demonstrated that visualization of English intonation led to improvement. Similarly, our Turkish L2 learners, without the use of technology-created visualizations, improved their production of one novel contour (i.e. falling–rising), and improved their perception for all contours. In other words, there is no reason to think that teaching intonation is any different from other phonological features. If we teach it, we should expect our learners to improve (Lee et al., 2015).

In pronunciation teaching, intonation is considered a difficult topic for teachers, who find it confusing and complicated, which may result in a loss of confidence for teachers and reduced willingness to teach a feature that confuses them (Couper, 2016; Kochem,

2022). Part of the confusion with intonation is that it is a feature that does not lend itself to judgments of right or wrong, as do phonemes and word stress. Teachers who desire unambiguous rules may therefore find intonation confusing (Kochem, 2022). This also means that intonation will not change the meanings of words and is highly unlikely to affect intelligibility in the ways measured by Jenkins (2000), who argued that intonation was not worth teaching. Instead, intonation is more likely to affect a different type of intelligibility, that is, the listener's understanding of a speaker's overall message or intentions (Levis, 2018). Applied to our study, teaching both the perception and production of rising, falling, and falling–rising intonation along with their functions is likely to succeed, but greater improvement may require longer treatments (Lee et al., 2015) or greater attention to developing awareness of intonational meaning (Liu, et al., 2024).

VI Future research and limitations

Several future directions for this research are immediately evident. First, the learning of intonation provides opportunities to explore the perception-production links in pronunciation learning. Intonation systems, although an important aspect of most languages, are often not evident to L2 learners who ignore them while attending to segmental, lexical, and syntactic features of spoken language (Pennington & Ellis, 2000). Intonation systems are learned early by native speakers, often before they begin to produce words, and their form-function connections differ across languages (Cruz-Ferreira, 1987). This study provides evidence that these Turkish learners, despite the lack of input outside the classroom, improved in both perception and production, and a future study could be designed to explore how perception improvements correlate with production improvements and vice versa.

Another possible topic to explore is longer-term improvement of the falling–rising contour. The falling–rising contour was interesting to the learners because of its novelty. Accuracy in its production started low, increased significantly, but always lagged behind the production of both pre-test and post-test productions of the other contours. It would be valuable to see whether production accuracy continued to increase with a longer period of instruction. It is possible that once the novelty wears off, accuracy in production would also level off. It would also be valuable to carry out a similar teaching intervention before learners took part in a study abroad experience in an English-speaking country (for word stress, see also the study of Romanelli et al., 2015). Jenkins (2000) argues that the best that can be expected of classroom instruction for many suprasegmental aspects of pronunciation (including intonation) is to raise awareness of their forms and functions. When learners later find themselves in a context where English is used widely, this awareness may help them to learn from the naturalistic input surrounding them.

Our study had limitations, including the smaller number of contours for rising intonation in comparison to those used for falling and falling–rising. Our inclusion of context did not include an equivalent number of instances for rising intonation as it did to the other contours, and this led to an unbalanced number of examples for the three contours.

A challenge in rating intonation production was classifying different intonation contours. Many learners produced very small rises and falls of pitch, so that we struggled to

unambiguously classify their productions. Most of our disagreements in classifying learners' productions occurred on these uncertain productions. It would be valuable to explore the differences in production of rising and falling intonation in their spontaneous spoken English and Turkish to help understand how they used pitch movement when they were not reading.

Our study shows that intonation, like other aspects of pronunciation, can be successfully taught, and that limiting instruction to a few important contours can be successful in helping learners perceive the contours and produce a subset of the contours. The study also demonstrates that our inclusion of context was a good start but needs to be more carefully operationalized. Indeed, future research on intonation learning should continue to explore context's role in learning intonation.

Acknowledgements

We would like to thank all learners in this project who were willing to share their classroom data to be used in this study. Additionally, we would like to express our deepest gratitude and appreciation to Dr Eryn Blagg for her invaluable consultation during our statistical analyses and Dr Pavel Trofimovich for providing feedback on an earlier version of this article.

Data Availability

The corresponding author will provide access to the data and R Markdown file if the reader needs online access.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Sinem Sonsaat-Hegelheimer https://orcid.org/0000-0002-2116-5213

References

Allen, V.F. (1971). Teaching intonation, from theory to practice. *TESOL Quarterly*, *5*, 73–81.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.

Baumann, S., & Grice, M. (2006). The intonation of accessibility. *Journal of Pragmatics*, *38*, 1636–1657.

Boersma, P., & Weenink, D. (2024). *Praat: Doing phonetics by computer* [computer program]. Available at: http://www.praat.org (accessed January 2025).

Bolinger, D. (1986). *Intonation and its parts: Melody in spoken English*. Stanford University Press.

Brauer, M., & Curtin, J.J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, *23*, 389–411.

Brazil, D. (1997). *The communicative value of intonation in English*. Cambridge University Press.

Brazil, D., Coulthard, M., & Johns, A. (1980). *Discourse intonation and language teaching*. Longman.

Burri, M., Baker, A., & Chen, H. (2017). 'I feel like having a nervous breakdown': Pre-service and in-service teachers' developing beliefs and knowledge about pronunciation instruction. *Journal of Second Language Pronunciation*, *3*, 109–135.

Chun, D.M. (1988). The neglected role of intonation in communicative competence and proficiency. *The Modern Language Journal*, *72*, 295–303.

Couper, G. (2016). Teacher cognition of pronunciation teaching amongst English language teachers in Uruguay. *Journal of Second Language Pronunciation*, *2*, 29–55.

Couper, G. (2019). Teachers' cognitions of corrective feedback on pronunciation: Their beliefs, perceptions and practices. *System*, *84*, 41–52.

Cruz-Ferreira, M. (1987). Non-native interpretive strategies for intonational meaning: An experimental study. In A. James & J. Leather (Eds.), *Sound patterns in second language acquisition* (pp. 103–120). Foris.

Cummins, C., & Rohde, H. (2015). Evoking context with contrastive stress: Effects on pragmatic enrichment. *Frontiers in Psychology*, *6*, Article 1779.

Dalton, C., & Seidlhofer, B. (1994). *Pronunciation*. Oxford University Press.

de Bot, K., & Mailfert, K. (1982). The teaching of intonation: Fundamental research and classroom applications. *TESOL Quarterly*, *16*, 71–77.

Derwing, T.M., & Munro, M.J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins.

Firke, S. (2023). *janitor: Simple tools for examining and cleaning dirty data: R package version 2.2.0* [software]. Available at: https://CRAN.R-project.org/package=janitor (accessed January 2025).

Fries, C.C. (1964). On the intonation of 'yes–no' questions in English. In D. Abercrombie, D. Fry, P. McCarthy, N. Scotte, & J. Trim (Eds.), *In honour of Daniel Jones: Papers contributed on the occasion of his eightieth birthday* (pp. 242–254). Longmans.

Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2022). *Package 'irr': Various coefficients of interrater reliability and agreement* [software]. Available at: https://cran.r-project.org/web/packages/irr/irr.pdf (accessed January 2025).

Geluykens, R. (1988). On the myth of rising intonation in polar questions. *Journal of Pragmatics*, *12*, 467–485.

Goh, C.C.M. (1994). Exploring the teaching of discourse intonation. *RELC Journal*, *25*, 77–98.

Gordon, J., & Darcy, I. (2016). The development of comprehensible speech in L2 learners: A classroom study on the effects of short-term pronunciation instruction. *Journal of Second Language Pronunciation*, *2*, 56–92.

Grant, L. (2016). *Well said* (4th ed.). Cengage.

Halliday, M.A.K. (1967). *Intonation and grammar in British English*. De Gruyter Mouton.

Hardison, D.M. (2004). Generalization of computer assisted prosody training: Quantitative and qualitative findings. *Language Learning and Technology*, *8*, 34–52.

Hartig, F. (2022). *DHARMa: Residual diagnostics for hierarchical (Multi-level / Mixed) regression models* [software]. Available at: https://CRAN.R-project.org/package=DHARMa (accessed January 2025).

Hedberg, N., Sosa, J.M., & Görgülü, E. (2017). The meaning of intonation in yes–no questions in American English: A corpus study. *Corpus Linguistics and Linguistic Theory*, *13*, 321–368.

Herment, S., & Tortel, A. (2021). The intonation contour of non-finality revisited. In A. Kirkova-Naskova, J. Fouz-González, & A. Henderson (Eds.), *English pronunciation instruction: Research-based insights* (pp. 175–194). John Benjamins.

Hewings, M. (2018). *English pronunciation in use (Advanced)*. Cambridge University Press.

Hewings, M., & Goldstein, S. (1999). *Pronunciation plus*. Cambridge University Press.

Hori, T., Toyama, M., & Akatsuka, M. (2020). Perception of English intonation by Japanese learners of English. In *Proceedings of the 10th International Conference on Speech Prosody 2020*, Tokyo, Japan, 25–28 May (pp. 838–842). Speech Prosody.

Hori, T., Toyama, M., & Akatsuka, M. (2021). Perception of English intonation patterns by Japanese learners of English. *Language Education and Technology*, *58*, 1–16.

Ipek, C. (2015). The phonology and phonetics of Turkish intonation. Doctoral dissertation, University of Southern California, Los Angeles, CA, USA. Publication number 11012093. ProQuest Dissertations Publishing.

Ipek, C., & Jun, S.-A. (2013). Towards a model of intonational phonology of Turkish: Neutral intonation. *Proceedings of Meetings on Acoustics*, *19*, Article 060230.

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford University Press.

Jiang, Y., & Chun, D. (2023). Web-based intonation training helps improve ESL and EFL Chinese students' oral speech. *Computer Assisted Language Learning*, *36*, 457–485.

Kenworthy, J. (1987). *Teaching English pronunciation*. Longmans.

Kabak, B. (2016). Refin(d)ing Turkish stress as a multifaceted phenomenon. *Paper presented at the Second Conference on Central Asian Languages and Linguistics, Indiana University*, 7–9 October.

Kochem, T. (2022). Second language teacher cognition development in an online L2 pronunciation pedagogy course: A quasi-experimental study. Doctoral dissertation, Iowa State University, Ames, IA, USA.

Ladd, D.R. (1996). *Intonational phonology*. Cambridge University Press.

Lane, L. (2012). *Focus on pronunciation 3* (3rd ed.). Pearson Education.

LaScotte, D., Meyers, C., & Tarone, E. (2022). *Voice and mirroring in L2 pronunciation instruction*. Equinox.

Le, H.T., & Brook, J. (2011). Using Praat to teach intonation to ESL students. *Hawaii Pacific University TESOL Working Paper Series*, *9*, 2–15.

Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics*, *36*, 345–366.

Lenth, R. (2023). *emmeans: Estimated marginal means, aka least-squares means: R package version 1.8.7* [software]. Available at: https://CRAN.R-project.org/package=emmeans (accessed January 2025).

Levis, J.M. (1997). Grammar and intonation: An alternative approach. *Speak Out*, *21*, 26–29.

Levis, J.M. (1999). Intonation in theory and practice, revisited. *TESOL Quarterly*, *33*, 37–63.

Levis, J.M. (2018). *Intelligibility, oral communication, and the teaching of pronunciation*. Cambridge University Press.

Levis, J.M., & Muller Levis, G. (n.d.). Pronunciation for a purpose. Unpublished book.

Levis, J.M., & Pickering, L. (2004). Teaching intonation in discourse using speech visualization technology. *System*, *32*, 505–524.

Levis, J.M., & Wichmann, A. (2015). English intonation–Form and meaning. In M. Reed & J. Levis (Eds.), *The handbook of English pronunciation* (pp. 139–155). Wiley Blackwell.

Liu, D., McGregor, A., Zielinski, B., Reed, M., & Meyers, C. (2024). ESL teachers' metalanguage as evidence of their metalinguistic knowledge of the English intonation system. *Language Awareness*. Advance online publication. https://doi.org/10.1080/09658416.2024.2370886

Lüdecke, D. (2023). *sjPlot: Data visualization for statistics in social science: R package version 2.8.15* [software]. Available at: https://CRAN.R-project.org/package=sjPlot (accessed January 2025).

Niebuhr, O., Alm, M., Schümchen, N., & Fischer, K. (2017). Comparing visualization techniques for learning second language prosody: First results. *International Journal of Learner Corpus Research*, *3*, 250–277.

Nolan, F. (2022). The rise and fall of the British school of intonation analysis. In J. Barnes & S. Shattuck-Hufnagel (Eds.), *Prosodic theory and practice* (pp. 319–361). The MIT Press.

O'Connor, J.D., & Arnold, G.F. (1961). *Intonation of colloquial English*. Longmans.

Ortega-Llebaria, M., & Colantoni, L. (2014). L2 English intonation: Relations between form-meaning associations, access to meaning, and L1 transfer. *Studies in Second Language Acquisition*, *36*, 331–353.

Pennington, M.C., & Ellis, N.C. (2000). Cantonese speakers' memory for English sentences with prosodic cues. *The Modern Language Journal*, *84*, 372–389.

Pickering, L. (2018). *Discourse intonation: A discourse-pragmatic approach to teaching the pronunciation of English*. University of Michigan Press.

Pike, K.L. (1945). *The intonation of American English*. University of Michigan Press.

R Core Team. (2023). *R: A language and environment for statistical computing* [software]. R Foundation for Statistical Computing. Available at: https://www.R-project.org (accessed January 2025).

Romanelli, S., Menegotto, A.C., & Smyth, R. (2015). Stress perception: Effects of training and a study abroad program for L1 English late learners of Spanish. *Journal of Second Language Pronunciation*, *1*, 181–210.

Shepherd, M.A. (2011). Functional significance of rising-intonation declaratives in settings with special discursive norms. *LSA Annual Meeting Extended Abstracts*, *2*, 10–11.

Sonsaat, S. (2017). The influence of an online pronunciation teacher's manual on teachers' cognitions. Unpublished doctoral dissertation, Iowa State University, Ames, IA, USA.

Tanner, M.W., & Landon, M.M. (2009). The effects of computer-assisted pronunciation readings on ESL learners' use of pausing, stress, intonation, and overall comprehensibility. *Language Learning and Technology*, *13*, 51–65.

Thompson, S. (1995). Teaching intonation on questions. *ELT Journal*, *49*, 235–243.

Ward, G.L., & Hirschberg, J. (1986). Reconciling uncertainty with incredulity: A unified account of the L*+ HLH% intonational contour. Paper presented at the Annual Meeting of the Linguistic Society of America. Available at: https://eric.ed.gov/?id=ED285390 (accessed January 2025).

Warren, P. (2016). *Uptalk: The phenomenon of rising intonation*. Cambridge University Press.

Wickham, H., & Grolemund, G. (2017). *R for data science: Visualize, model, transform, tidy, and import data*. O'Reilly Media. Available at: https://r4ds.had.co.nz (accessed January 2025).

Wickmann, A. (2015). Functions of intonation in discourse. In M. Reed & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 139–156). West Sussex, UK: Wiley-Blackwell.

Winter, B. (2020). *Statistics for linguists: An introduction using R*. Routledge.

Wong, R. (1987). *Teaching pronunciation: Focus on English rhythm and intonation: Volume 68*. Center for Applied Linguistics, Washington DC, USA.

Zhang, R., & Yuan, Z.M. (2020). Examining the effects of explicit pronunciation instruction on the development of L2 pronunciation. *Studies in Second Language Acquisition*, *42*, 905–918.

Appendix A

Intonation topics and exercises in class

	Forms and functions of intonation
Week 12	• Importance of intonation in meaning • Finished and unfinished statements (Fall or Rise) • Certainty, uncertainty, surprise (Fall or Rise) • *Wh*-questions seeking for information or repetition (Fall or Rise) • Yes/no questions (Rise)
Types of Exercises	• Identify the intonation patterns in the conversation (Does the speaker's voice fall or rise?) • Read the conversation with the same intonation patterns. • Listen and repeat the questions. • Listen to the conversation and identify the situation between speakers • Identify if the speaker in each line of the conversation has finished speaking. • Shorten long sentences by keeping the meaning through intonation patterns. • Put the mixed-up conversations in the correct order and mark each line of the conversation as rising or falling according to meaning • Work with a partner. Student A reads a statement. Student B responds with a *wh*-question using either falling or rising intonation. Student A responds with the matching answer.
Week 13	• Affirmative statements that are finished (Fall) • Seeking out information questions (Fall) • Repetition or surprise questions (Rise) • Yes/no questions (Rise) • Stating reservation, reluctance, etc. (Fall–Rise) • Questions of choice with 'Or' (Rise + Fall) • Listing a series of items (Rise, Rise, Fall)
Types of Exercise	• Identify the intonation pattern at the end of the statement (Fall or Rise) • Identify the intonation pattern of the question (Fall or Rise) • Identify the intonation pattern of each statement in the dialogue (Fall or Rise) • Listen to each phrase in the short dialogue and identify the intended meaning • Identify the new intonation pattern you hear. What does it sound like? (Fall–Rise) • Identify if the intonation patterns of the two phrases is the same or different (Fall – Fall–Rise) • Identify the intonation patterns in paragraph (a TED talk 'What causes insomnia?') • Listen to the dialogue between two speakers and identify the intended meaning of Speaker B (reservation or reluctance) • Say each of the given phrases with Fall, Rise, or Fall–Rise intonation • Read Speaker A and B's statements in the conversation based on the meaning • Read the statements in a way to show an alternative option between two things. • Listen and repeat the anecdotes in chorus

(Continued)

Appendix A. (Continued)

	Forms and functions of intonation
Week 14	• Finished statements (Fall) • Seeking out for information *wh*-questions (Fall) • Repeated questions (negotiation of meaning) (Rise) • Unfinished statements (Rise) • Questions of choice with 'Or' (Rise + Fall) • Listing a series of items (Rise, Rise, Fall) • Reservation, reluctance, etc.
Types of Exercises	• Identify the intonation patterns of phrases and sentences • Deliver a 1-minute talk where you talk about one of your hobbies in class • Monitor your classmate's intonation patterns to provide feedback on the patterns your classmate used and their connection to their intended meaning.

Appendix B

Perception and production test items (identical sets).

Part 1: Forms of intonation patterns

Instruction for perception test: Identify the intonation contour of the utterance you hear. (Falling, Rising, Falling–Rising)

Instruction for production test: Read the utterances with the given intonation contour.

1.	Ready	Falling
2.	Maybe	Rising
3.	Hungry	Rising
4.	Almost	Falling–Rising
5.	Tired	Falling
6.	Bored	Falling
7.	Happy	Falling–Rising
8.	Sure	Rising
9.	Ok	Falling
10.	Perhaps	Rising
11.	Soon	Falling
12.	Excited	Falling–Rising
13.	Upset	Rising
14.	Sort of	Falling–Rising
15.	Like it	Falling–Rising

Part 2: Functions of intonation patterns

Instruction for perception test: Listen to the utterance and choose the intended meaning of the speaker.

Instruction for production test: Read the utterances with the appropriate intonation contour to give the intended meaning.

1.	I'd love to	Falling	(And I will.)
2.	It's ok	Falling	(It is really okay.)
3.	Looks interesting	Falling–Rising	(But still not very promising)
4.	I'd love to	Falling–Rising	(But I cannot)
5.	It's ok	Falling–Rising	(It's not actually okay.)
6.	I think I could	Falling	(I will be able to.)
7.	Funny guy	Falling	(He's a funny guy.)
8.	Looks interesting	Falling	(It is interesting indeed.)
9.	I think I could	Falling–Rising	(But I need to check my schedule.)
10.	Would be helpful	Falling	(It will help.)
11.	Not a good show	Falling	(Don't waste your time with it.)
12.	Would be helpful	Falling–Rising	(It's helpful but not sufficient.)
13.	Not my first choice	Falling–Rising	(But it still is not bad.)
14.	Funny guy	Falling–Rising	(Funny guy but not my type.)
15.	Not a good show	Falling–Rising	(But it still keeps me busy.)
16.	Not my first choice	Falling	(This wasn't the one I chose best.)

Part 3: Functions of intonation patterns

Instruction for perception test: Listen to the short dialogue between Speaker A and B. Choose the intended meaning conveyed by Speaker B.

Instruction for production test: Read the Speaker B's response to reflect the intended meaning.

Notes. F: Falling; FR: Falling–Rising

1.	Speaker A:	Hey, can you help me?	
	Speaker B:	I'd be glad to help you.	(I'd love to but I really don't have time. / FR)
2.	Speaker A:	Do you think this will make the boss happy?	
	Speaker B:	There's a lot that has been done.	(A lot has been done, but not enough. / FR)
3.	Speaker A:	Hey, can you help me?	
	Speaker B:	I'd be glad to help you.	(I will gladly help with this. / F)
4.	Speaker A:	I'll have to miss three class sessions this semester. Do you think the professor will be fine with it?	
	Speaker B:	Oh, I'm sure it's fine.	(Yeah, I don't think she cares about attendance. / F)

5.	Speaker A:	Do you think I should print the assignment on campus?	
	Speaker B:	We've a printer in the dorm.	(But it prints only in black and white. / FR)
6.	Speaker A:	Do you think this will make the boss happy?	
	Speaker B:	There's a lot that has been done.	(We have done a lot. She'll like it. / F)
7.	Speaker:	I'll have to miss three class sessions this semester. Do you think the professor will be fine with it?	
	Speaker B:	Oh, I'm sure it's fine.	(I think it's fine but it's still good to let her know in advance. / FR)
8.	Speaker A:	How do you like your teaching schedule?	
	Speaker B:	It's ok.	(It's fine but could be better. / FR)
9.	Speaker A:	Do you think I should print the assignment on campus?	
	Speaker B:	We've a printer in the dorm.	(No need. We have a printer in the dorm. / F)
10.	Speaker A:	You think the test was hard?	
	Speaker B:	Sort of.	(It was but not something I couldn't do. / FR)
11.	Speaker A:	How do you like your teaching schedule?	
	Speaker B:	It's ok.	(I think it's fine. / F)
12.	Speaker A:	Speaker A: You think the test was hard?	
	Speaker B:	Sort of.	(Yes, it really was hard. / F)