

English Intonation Perception and Production: A Systematic Review

LU MINGLONG

*Faculty of Social Sciences and Humanities,
Universiti Kebangsaan Malaysia, Malaysia*

SHARIFAH RAIHAN BINTI SYED JAAFAR *

*Faculty of Social Sciences and Humanities,
Universiti Kebangsaan Malaysia, Malaysia
s_raihan@ukm.edu.my*

ABSTRACT

Previous English intonation studies concentrate too much on intonation production though some of them have already investigated intonation from the perception perspective. The relationship between the perception and production of English intonation is still not clear and there are conflicting views. The present study follows PRISMA protocol and identifies collectively 19 studies from WOS, Scopus and Proquest to include as many papers as possible in the research pool. Through generating themes and codes, this review has compared their methodologies, synthesised their findings from “Tonality, Tonicity and Tone” (chunking, focus and tunes). Mixed-mode designs are the most popular design methods in researching Tonicity and Tone but Tonality has drawn less attention. Participants from tertiary education are frequently chosen while intervention and measurement validity are less mentioned. The perception-production relationship of English intonation is either asymmetric or correlational. Besides, this review has summarised the directions for the future English intonation research concerning the perception-production relationships. Future research should concentrate on various phonetic dimensions, their functions in production-perception relationship of intonation, discrepancy of certain acoustic cues in different sentence types, more diversified language varieties, pedagogical and pragmatic approaches. Most importantly, in the intonation teaching field, future research should explore different L2 speakers, contextual information, retention of intonation acquisition and L1 transfer. It is also of great significance to further clarify the perception-production relationships.

Keywords: Perception; Production; English Intonation; prosody; suprasegmentals

INTRODUCTION

Intonation is an important feature of speech for some intonational languages including English, Spanish, German and so on. However, there is still no commonly agreed definition of this significant element of spoken language. Cruttenden (1997) introduced intonation from the forms, which include boundary, nucleus and pitch contours. Vaissière (2005) remarked especially from the perception perspective that intonation is strictly restricted to the perceived F0 pattern, or the perception of other prosodic parameters fulfilling the same functions: pauses, relative loudness, voice quality, duration, and segmental phenomena. Levis (2012) maintained that intonation refers to a combination of acoustic parameters, including duration, intensity, and pitch used to communicate discourse meaning, emphasizing the purposive discourse intentions. Generally speaking, a narrow description of intonation would only cover pitch movements or tones (tunes) while this paper discusses English intonation from a more inclusive view which involves Tonality, Tonicity and Tone (Wells, 2006). This definition from Wells is chosen because it comprehensively covers all the major components most often discussed across different definitions and theories. Tonality means the chunking process of an utterance into several intonation phrases that contain

information the speaker wants to convey. Tonicity, sometimes also referred to as sentence stress or information focus, stands for the placement of tonic syllable or the nucleus. Tone or nuclear tone can be defined as the specification of the contrastive pitch contours on the tonic syllable and the following syllables in the tail.

On the one hand, the Production of English intonation has long been the most popular perspective that numerous studies mainly focus on. For example, different teaching methods are utilised to help produce better English intonation, including visual displays (Zhi et al., 2024), video-based shadowing (Van et al., 2024), re-synthesised self-produced stimuli (Li et al., 2020), advertising (Demina, 2023), multi-modal methods (Zhao et al., 2020), and so on. Past studies predominantly employ various teaching methods and evaluate their effects on intonation production. In addition, they also overwhelmingly follow the production-oriented perspective. Varied factors are taken into account to analyse intonation production or acquisition, ranging from L1 transfer (Kang, 2024) to cross-linguistic influence (Li & Feng, 2024), English proficiency (Tian et al., 2024), proficiency and study-abroad experience (Kang et al., 2021), acoustic and eye-tracking experiments (Liu & Reed, 2021) and so forth. On the other hand, some recent studies have already extended the focus from intonation production to intonation perception. For instance, visual cues and facial expression (Li et al., 2024), gestures (Hori et al., 2025), durational cues (Hori et al., 2020) and contextual information (Radu et al., 2020) are applied to the discussion of perceiving English intonation, especially pitch contours. Recently, the perception-centred perspective is gaining momentum in English intonation studies. The perception and production of L2 prosody are often analysed separately, but the link between the two is rarely investigated (Puga et al., 2018; Ortega-Llebaria & Colantoni, 2014; Altmann, 2006).

As it is mentioned above, previous studies are predominantly conducted on teaching methods and English intonation production while some research would alternatively cover English intonation perception. However, only a few studies are aimed at giving a comprehensive analysis and description of all the factors including perception, intonation training, production and their connections. Besides, conflicting opinions show that the perception-production relations are either asymmetric (Riquelme et al., 2025; Sonsaat-Hegelheimer & Levis, 2025; Hudson et al., 2022; Cao & Li, 2022) or correlational (Zhao et al., 2022). Thus, the topic of intonation perception-production relationships is under-explored and thus unclear. Figure 1 depicts the perception-production relationships of English Intonation. To investigate this specific research area can help us better understand the relations between these factors and what's happening at each stage (such as perception, teaching, production) during the experiments in different studies, which lays the very foundation for future related research topics into the intonation acquisition of tonal languages and non-intonation languages alike.

Considering the predominance on English intonation production and seemingly complicated perception-production link (Puga et al., 2018), it is imperative to conduct more studies to help shed light into this specific research domain of perception-production relationships. Unfortunately, no systematic review has encompassed studies that have been conducted to delve into the perception and production of English intonation, so as to provide a comprehensive synthesis of their relationships and map out the research scope. Consequently, there is a pressing need for the composition of a comprehensive systematic literature review to resolve these discrepancies, evaluate methodologies, compare research contexts and search for directions for English intonation perception and production studies in the future.

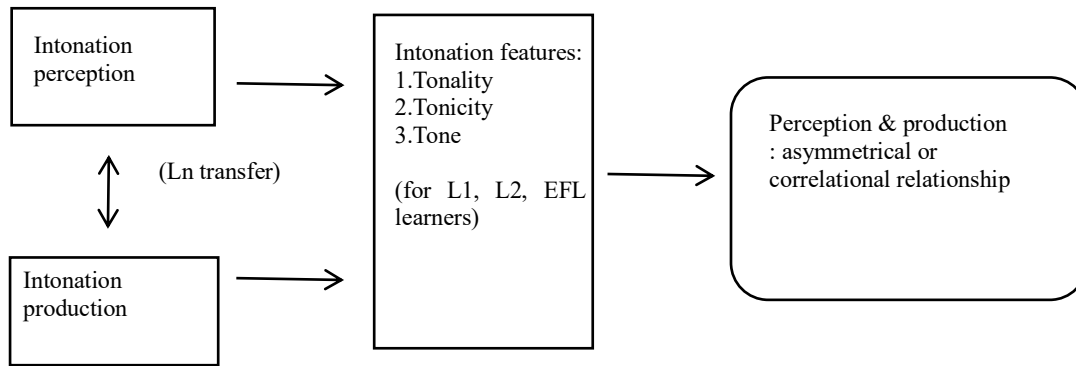


FIGURE 1. The Perception-Production Relationships of English Intonation

THE PRESENT STUDY

The present study bridges the aforesaid gap by conducting a systematic literature review (SLR) of empirical studies and doctoral dissertations from inception to 2025 on the perception and production of English intonation in WOS (Web of Science), Scopus and Proquest. Doctoral dissertations are also included because this research area has not received due attention and there are not so many published articles for reference. Besides, doctoral dissertations are also of high quality and worthy of being referenced.

The present study intends to achieve the following research objectives:

1. to compare existing research methodology of English intonation perception and production, including research designs, participant profiles, instruments, intervention;
2. to clarify the distribution of research domains and language pairing;
3. to summarise findings of the perception-production relationships from previous studies;
4. to search for research limitations, gaps and suggest directions for future study.

The present study is guided by the following research questions:

1. How does existing English intonation perception production research differ in methodology, in terms of research designs, participants, instruments, intervention?
2. What is the distribution of research domains and language pairing?
3. What are the findings for past studies on English intonation perception and production?
4. What methodological limitations and research gaps can be concluded through analysing previous studies?

METHODOLOGY

In order to guarantee that the methodology is reliable and replicable, this review follows PRISMA (the Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines 2020 (Page et al., 2021). The search for past studies included published or online empirical studies and indexed doctoral dissertations from inception to 2025 in WOS, Scopus and Proquest. The purpose of

leaving the start date open was to capture as many search results as possible, showcasing theoretical and methodological development in the study of English intonation perception and production. The purposive focus on this period helps to produce a more comprehensive synthesis of past studies and better reflects their conceptual frameworks, research designs, instruments, and pedagogical implications.

DATABASES AND SEARCH STRATEGY

The selection standard for databases was driven by the demand for well-rounded, good-quality coverage of language studies and linguistics. Web of Science (core collection) was selected for its wide index of prestigious, cross-disciplinary research, which enables access to frequently cited high-impact studies in the relevant area. Scopus was chosen because it is one of the biggest abstract and citation databases of peer-reviewed literature including scientific journals, books and conference proceedings. ProQuest was also included because of its extensive collection of doctoral dissertations and peer-reviewed journal articles. ERIC was not chosen mainly because its coverage is partly overlapped with the three databases mentioned above.

The search strings in the three databases included combinations of the following keywords using Boolean operators and Wildcards: (perce* and produc* English intonation*) OR (perce* and produc* English nuclear stress) OR (perce* and produc* English nuclear accent*) OR (perce* and produc* English prominence) OR (perce* and produc* English nuclear tone*) OR (perce* and produc* English nuclear tune*) OR (perce* and produc* English phras*) OR (perce* and produc* English paus*) OR (perce* and produc* English chunking). The search strings in the three databases need to include all the possible elements of English intonation including Tonality, Tonicity and Tone, so different expressions are used to ensure the precision and inclusion of targeted search of literature.

INCLUSION AND EXCLUSION STANDARD

To guarantee the relevance and quality of selected articles, the following inclusion and exclusion standard was employed, as shown in Table 1:

TABLE 1. Inclusion and exclusion standard

Criteria	Inclusion standard	Exclusion standard
Scope	Studies (include doctoral dissertations) with explicit focus on both the perception and production perspectives	Studies (include doctoral dissertations) addressing only perception or only production
Field	Studies (include doctoral dissertations) in the field of English intonation, prosody and suprasegmentals	Studies (include doctoral dissertations) in fields like translation, literature, phonology, segmentals
Type	Empirical studies; research articles; doctoral dissertations	Theoretical articles; editorials, books or review articles; master's thesis
Population	ordinary L1, L2, EFL learners	autistic, hearing-impaired, speech-impaired populations
Time span	2025 and prior	after 2025

SCREENING AND SELECTION PROCESS

Following PRISMA's procedures (Identification, Screening, Included), an initial pool of 1345 articles was retrieved from the three academic databases: Web of Science (n=524, after copying the search strings directly to the search column of Web of Science and controlling the time span), Scopus (n=475, after searching within "Article title, Abstract, Keywords" and controlling the time span) and ProQuest (n=346, after advanced searching within "Title OR Abstract" and controlling the time span). All the data were downloaded from the three databases as RIS format files to be uploaded to software Zotero. The software Zotero 9 was equipped with the Plugin "Zoplicate" (Version 5.0.9 authored by Chenglong Ma) to help detect and manage duplicate items. Data from Web of science were uploaded first, followed by data from Scopus and Proquest. When removing duplicates, the earliest record would be kept if an article was indexed in more than one database. After removing duplicates by using Zotero 9 software, 1024 different records remained. These records were then screened by title and abstract, reducing the pool to a smaller subset for full-text review. When screening the articles, the "Note" function of the software can be used to create notes based on the title, abstract and the full text. The notes written can be searched later for navigating and narrowing the target pool of studies. For instance, the self-created note by the author "per into tone sentence type PhD" means "perception of intonation; tone; sentence types; PhD dissertation". These notes can not only help classify the studies by adding different marks but also help quickly retrieve them by using the Advanced Search Function. The screening and selection process and the checking process were completed by the first author. A total of 72 studies were marked with reminding notes and then examined in full, from which 19 studies met all the inclusion criteria and thus were included in the final synthesis. A diagram of the selection process is presented in Figure 2.

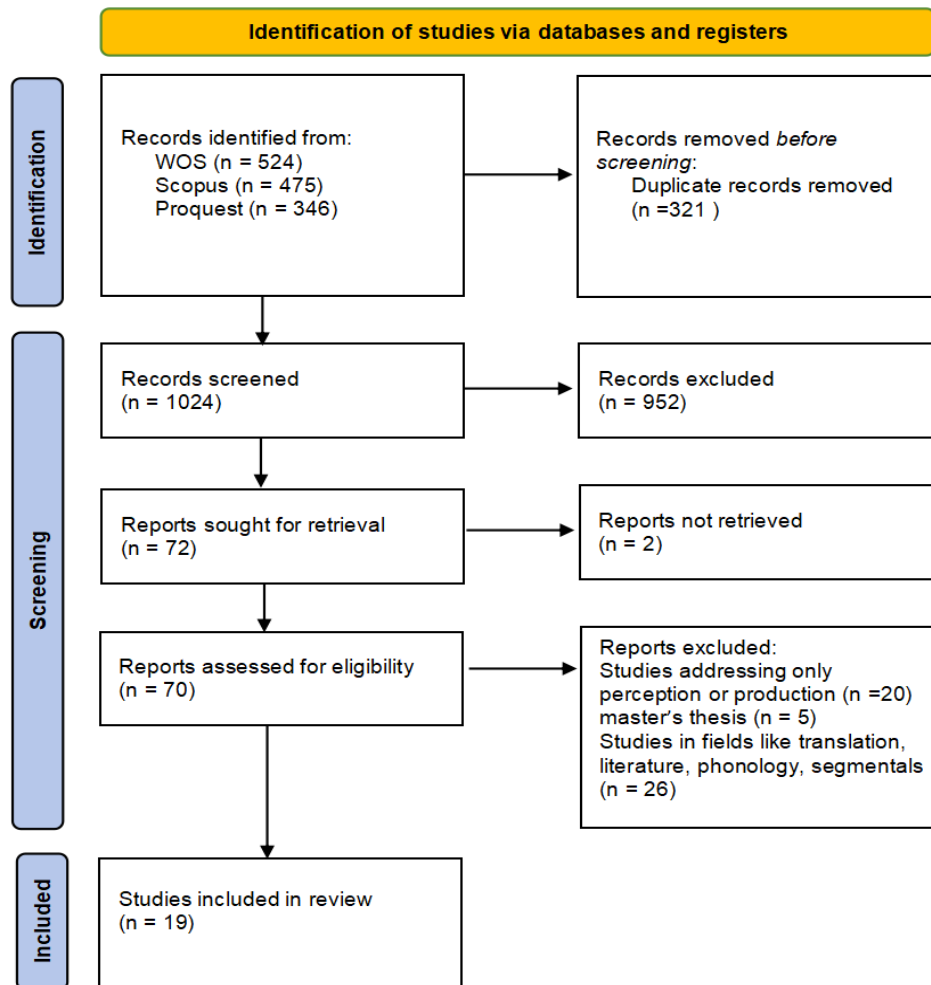


FIGURE 2. A diagram of the screening and selection process

DATA COLLECTION AND ANALYSIS

The data extraction and coding were conducted by the first author and coded themes were examined by an independent reviewer to improve reliability. Discrepancy was resolved by discussions between them. Data from each selected studies were extracted by using Excel, which recorded data categories including metadata (author names and publication year), the context (language paring and research domains), methodology and qualitative text findings (results, frameworks, limitations and future directions). The thematic content analysis approach (Ahmed et al., 2025) was chosen to analyse the included studies. First, to read and classify the studies to get familiar and immersed. Identify initial ideas, patterns and concepts. Detect prominent features and label them to generate codes for analysis, and the coding process was performed and categorised according to the research objectives and research questions. After that, the codes were grouped into meaningful themes. These themes were reviewed to guarantee analytical rigor, clarity, and consistency with the research objectives and research questions. Finally, the synthesised results were reported in the following section according to the research objectives, research questions and the themes decided.

RESULTS

Results for the present study are presented in the following sections from: research designs, participants, statistical methods, instruments and intervention; research domains and language pairing; the findings of perception-production relationship; methodological limitations and future research directions in English intonation perception and production.

RESEARCH DESIGNS, PARTICIPANTS, STATISTICAL METHODS, INSTRUMENTS AND INTERVENTION

As can be seen from Table 2, for the research design, five studies (Hudson et al., 2022; Sonsaat-Hegelheimer & Levis, 2025; Riquelme et al., 2025; Katsika et al., 2023; Wang et al., 2010) were mainly quantitative, with their data collected, quantified and measured statistically. However, the majority of the selected studies employed a more favourable and popular mixed-mode research design, enjoying the advantages of both quantitative and qualitative methods in analysing production and perception data.

The education background and proficiency of participants profile are given more attention than the number and age of them because the latter two characteristics (number and age) of the sample vary to a large extent and it will be less meaningful to report this information. The vast majority of the studies selected university students, among which some participants were English majors or with linguistic backgrounds. This is mainly because the researchers for these studies are from academic circuit and they just conduct purposive sampling for availability or convenience. Surprisingly, one study involved participants from secondary school (Colantoni, et al., 2022), extending the research population to cover different education levels and stages.

In contrast, there is a wide spectrum of different proficiency levels in the research participants. Four studies sampled a more diversified participants of different proficiency levels (Riquelme et al., 2025; Swerts & Zerbian, 2010; Buzan et al., 2022; Wang et al., 2010). Six studies only selected participants of advanced level while two studies chose participants of intermediate level (Zhao et al., 2022; Cao & Li, 2022). However, some studies did not state clearly the proficiency of their participants. Nevertheless, considering their identity as English majors, undergraduate or master students, it is evident that most of them can be categorised as intermediate or around. For researchers, participants of intermediate level are more approachable, and they also can represent the proficiency level of most L2 or EFL learners.

In the statistical analysis part, different methods were employed to decode the data collected. T-test and ANOVA were widely used across different studies, which are statistical tests used to determine if differences between group means are statistically significant. A t-test compares the means of two groups, while ANOVA (Analysis of Variance) compares the means of three or more groups. Another widely mentioned method was the post hoc tests that involve pairwise comparison procedures conducted after a significant Omnibus ANOVA result ($p < 0.05$). Post hoc tests are significant when an independent variable has three or more groups, because ANOVA only indicates that at least one group is different, but Post hoc tests can determine specifically which group means differ. Notably, Ortega-Llebaria and Colantoni (2014) took non-parametric tests, since not all the data are parametric. Therefore, sometimes, there is a need to seek for non-parametric tests to find out the target results.

Almost all the studies used similar production and perception tests as instruments. In production, speaking was recorded and software might be used to transcribe or investigate acoustic cues while in perception, there were listening to tasks (often forced-choice questions online or

offline). These are the most common and popular instruments in the production and perception of English intonation for quantitative data. On the other hand, qualitative data sometimes also included data from questionnaires, surveys and interviews (Gasser et al., 2019; Menjo, 2018; Katsika et al., 2023; Cho, 2018).

Intervention is not necessarily a must-have for all the studies. This is because some of them follow the control-experiment model while others adopt the pretest-posttest model. In the pool of included studies, only one researcher applied intonation training or teaching as a kind of intervention and compared the results before and after intervention (Sonsaat-Hegelheimer & Levis, 2025). Therefore, there is a great potential for developing and experimenting studies with intervention or doing pedagogical research in English intonation perception and production, due to the scarcity.

TABLE 2. Research designs, participants, statistical methods, instruments and intervention

Theme	Code	Studies
Research designs	Mainly quantitative	Hudson et al. (2022); Sonsaat-Hegelheimer & Levis (2025); Riquelme et al. (2025); Katsika et al. (2023); Wang et al. (2010)
	Mixed-mode	Colantoni et al. (2022); Zhao et al. (2022); Cao & Li (2022); Ulbrich (2010); Swerts & Zerbian (2010); Buzan et al. (2022); Ortega-Llebaria & Colantoni (2014); Stepling & Montgomery (2002); Gasser et al. (2019); Colantoni et al. (2016); Menjo (2018); Puga et al. (2018); Ali & Taha (2022); Cho (2018)
	Secondary	Colantoni et al. (2022)
Education (participants)	University Students	Colantoni, et al., (2022); Hudson et al. (2022); Zhao et al. (2022); Cao & Li (2022); Riquelme et al. (2025); Ulbrich (2010); Swerts & Zerbian (2010); Menjo (2018); Puga et al. (2018); Ali & Taha (2022); Cho (2018)
	Diverse	Riquelme et al. (2025); Swerts & Zerbian (2010); Buzan et al. (2022); Wang et al. (2010)
Proficiency (participants)	Advanced only	Hudson et al. (2022); Colantoni et al. (2022); Ortega-Llebaria & Colantoni (2014); Colantoni et al. (2016); Puga et al. (2018); Cho (2018)
	Intermediate only	Zhao et al. (2022); Cao & Li (2022);
	T-test	Riquelme et al. (2025); Buzan et al. (2022); Gasser et al. (2019); Ali & Taha (2022); Cho (2018)
Statistical methods	ANOVA	Ulbrich (2010); Ortega-Llebaria & Colantoni (2014); Stepling & Montgomery (2002); Menjo (2018); Cho (2018); Wang et al. (2010)
	Post hoc tests	Colantoni et al. (2022); Katsika et al. (2023); Ulbrich (2010); Ortega-Llebaria & Colantoni (2014); Menjo (2018); Puga et al. (2018); Wang et al. (2010)
	Other tests	Cho (2018) (Chi-squared Test) Ortega-Llebaria & Colantoni (2014) (non-parametric Kruskal-Wallis test)
	Quantitative	Almost all the studies: recording reading or speaking and analysing by Software or humans for production while administering listening tests (forced-choice) for perception
Instruments	Qualitative (Questionnaire s, surveys)	Gasser et al. (2019); Menjo (2018); Katsika, et al., (2023); Cho (2018)
	Applied	Sonsaat-Hegelheimer & Levis (2025)
Intervention	Not applied	The rest of the studies

RESEARCH DOMAINS AND LANGUAGE PAIRING

When it comes to the specific research domain, only three study (Menjo, 2018; Gasser et al., 2019; Cho, 2018) covered the area of Tonality. Tone (intonation patterns or tunes) is one of the most popular research areas in English intonation studies while Tonicity (nuclear stress or focus) is another popular domain. Sometimes, these two features are covered in one combined study.

As for which language background the participants were from or the language pairing, three studies focused on American English only while Chinese language (mainly Cantonese) seemed to account for a big proportion. Three studies researched Cantonese EFL learners (Hudson et al., 2022; Zhao et al., 2022; Cao & Li, 2022) and one study used Mandarin-English speakers as the control group (Ortega-Llebaria & Colantoni, 2014). However, only one research group (Wang et al., 2010) mainly experimented on Mandarin-English speakers. The rest of the studies exhibited a great variety of language pairing, which involved L2 or EFL speaker from Asia, Middle East or Europe, even with minorities from Africa (Zulu). Table 3 below provides a detailed description of the language background or language pairing information. However, the existing literature is fragmented and scattered, consolidating a research gap that is under-explored for the perception-production perspective of English intonation for L2 or EFL learners.

TABLE 3. Research domains and language pairing

Theme	Code	Studies
Research domains	Tonicity	Riquelme et al. (2025); Katsika et al. (2023); Ortega-Llebaria & Colantoni (2014); Wang et al. (2010)
	Tone	Hudson et al. (2022); Colantoni et al. (2022); Sonsaat-Hegelheimer & Levis (2025); Ulbrich (2010); Buzan et al. (2022); Stepling & Montgomery (2002); Colantoni et al. (2016);
	Tonicity and Tone	Zhao et al. (2022); Cao & Li (2022); Swerts & Zerbian (2010); Puga et al. (2018); Ali & Taha (2022);
	Others	Menjo (2018) (Tonality and Tonicity); Gasser et al. (2019) (Tonality, Tonicity, Tone); Cho (2018) (Tonality, Tone)
	American English only	Katsika et al. (2023); Stepling & Montgomery (2002); Gasser et al. (2019)
Language pairing ("*" here stands for English)	Cantonese	Hudson et al. (2022); Zhao et al. (2022); Cao & Li (2022)
	/*	Ortega-Llebaria & Colantoni (2014); Wang et al. (2010)
	(Mandarin/*)	
	Iraqi/*	Ali & Taha (2022)
	Inuktitut/*	Colantoni et al. (2022)
	Turkish/*	Sonsaat-Hegelheimer & Levis (2025)
	Chilean/*	Riquelme et al. (2025)
	Zulu/*	Swerts & Zerbian (2010)
	Portuguese/*	Buzan et al. (2022)
	Spanish/*	Ortega-Llebaria & Colantoni (2014); Colantoni, et al. (2016)
	Japanese/*	Menjo (2018)
	German/*	Ulbrich (2010); Puga et al. (2018)
	Korean/*	Cho (2018)

THE PERCEPTION-PRODUCTION FINDINGS OF ENGLISH INTONATION

This review also reports the findings from all the target studies. The results from previous related literature can help us draw on past studies and better understand the perception-production relationships in English intonation research. The results are discussed respectively from the three aspects of English intonation, including Tonality, Tonicity and Tone.

THE PERCEPTION AND PRODUCTION OF ENGLISH TONALITY

Out of all the studies, only Menjo (2018), Gasser et al. (2019) and Cho (2018) mentioned Tonality. This feature was presented by Menjo as silent pause and filled pause. Pauses are the major indicator to separate larger sentences into smaller intonation phrases, marking Tonality and tone boundaries. According to Menjo, as for silent pauses, the difference between native English speaker and 10-year-resident (Japanese) was not significant, so was the difference between 3-year-resident and non-resident participants. However, the difference between 10-year-resident and 3-year resident was significant. With regard to the number of filled pauses, the differences among all the speaker groups were not significant. Besides, the difference between non-resident participants and the other three speaker groups was significant in terms of silent pause duration. Gasser et al. (2019) reported that newscasters used significantly more intonation phrases than non-newscasters while there was no significant difference in the number of intermediate phrases between these two groups of participants. Cho (2018) stated that Korean speakers of English used a different parsing (chunking) pattern.

When it comes to the perception of Tonality, Menjo's study (2018) used a different expression, chunking. Native English speaker scored the highest and the mean scores of the Japanese speaker groups were similar to each other. The difference between native English speaker and the Japanese speaker groups was significant. To conclude, Menjo's study introduces the perception and production of English Tonality by comparing the performance between Japanese English speakers and native English speakers. Gasser et al. (2019) compared the number of intonation phrases and intermediate phrases between newscasters and non-newscasters. More research on this domain should be encouraged since few of them would cover the topic of English Tonality.

THE PERCEPTION AND PRODUCTION OF ENGLISH TONICITY

Cao and Li (2022) believed that there were no strong correlations between the production and perception of narrow focus for Cantonese English speakers. Similarly, Riquelme et al.(2025) discovered that working memory and English proficiency had no significant associations with the perception or production of nuclear accent and there was asymmetry between its perception and production. These findings were consistent with that of Ali and Taha (2022), who found out that Iraqi EFL learners recognised sentence stress better than produced it. However, Zhao et al., (2022) reported that for Cantonese English speakers, their perception and production accuracy of nuclear accent placement showed similar tendency (correlations): participants performed better in both perception and production when the nuclear position was initial and medial (compared with final and broad nucleus position). Wang et al. (2010) proved the close relationship between the Chinese EFL learners' oral English proficiency and their production and perception of prominence distribution patterns.

Puga et al. (2018) discussed some nucleus placement errors. Swerts and Zerbian (2010) emphasised L1's transfer to L2 and maintained that proficiency affected this transfer. Menjo (2018) discussed L1's transfer from the period of residence for participants. Ortega-Llebaria and Colantoni (2014) held that adding context increased L1 transfer effects on the perception and production of focus and that L1 transfer was more evident in production than in perception. To sum up, previous research findings can be quite paradoxical and confusing in that "no correlations/associations" coexists with "correlations" and "asymmetry" in the perception and production of English

Tonicity. Other perspectives that frequently occur in this research domain include L1 transfer (possibly both positive and negative) and the positive effect of intonation training or intervention.

THE PERCEPTION AND PRODUCTION OF ENGLISH TONE

Tone, sometimes referred to as tunes, intonation, pitch contour or pitch movement, is another very popular research domain in English intonation studies. Hudson et al. (2022) found out in their experiments that participants performed better at production of nuclear tones than in the corresponding recognition study. This result is supported by Puga et al. (2018). They discovered that German English speakers performed better in production than in perception task for nuclear tones. Ali and Taha (2022) also concluded that production of intonation preceded its recognition. Moreover, Colantoni et al. (2022) identified similar but more complex results. In their study, most participants had higher accuracy in production while some participants were the opposite or performed equally in perception and production. Sonsaat-Hegelheimer and Levis (2025) also detected asymmetry in perception and production of Turkish speakers of English in terms of tone. On the other hand, Zhao et al. (2022) noted that Cantonese EFL tended to use H* and L* in statements and yes-no questions respectively in production, and these two nuclear tone types were also the easiest to be perceived, suggesting there is a correlation.

Ulbrich (2010) argued that prosodic similarities between L1 and L2 facilitated the acquisition of L2 prosody, which implied positive L1 transfer in prosody. On the contrary, Cho (2018) reported that Korean EFL learners struggled to perceive and express English emotional intonation due to the negative transfer of L1 prosodic habits, such as flatter pitch contours. Swerts and Zerbian (2010) also emphasised L1's transfer to L2 and maintained that proficiency affected this transfer. Buzan et al. (2022) further supported L1 intonational interference in both production and perception of L2 utterances. Colantoni et al. (2016) believed that Spanish speakers of English demonstrated difficulty acquiring the distinction between absolute yes-no questions and declarative questions while context increased the degree of cross linguistic transfer. Here, Tone tends to be discussed with sentence types because tone characteristics can be distinct between different types of sentence types. Interestingly, Gasser et al. (2019) summarised that listeners could distinguish newscaster from conversational speech by prosody, emphasizing pragmatic and prosodic functions across registers and genres.

In conclusion, asymmetry between the perception and production of English tones can be found in many studies, and the production seems to be better than perception performance for participants in many cases. However, there is still little evidence that correlational interaction may also exist between the perception and production of English tones, especially for Cantonese speakers of English (Zhao et al., 2022). Other factors like L1 transfer and the positive training effects can also be spotted most of the time.

METHODOLOGICAL LIMITATIONS AND FUTURE RESEARCH DIRECTIONS IN ENGLISH INTONATION PERCEPTION AND PRODUCTION

For this part, limitations on measurement validity, research design and future research directions are mainly introduced in Table 4. Out of the total 19 studies, only 5 studies mentioned validity or reliability in their papers (Sonsaat-Hegelheimer & Levis, 2025; Ulbrich, 2010; Stepling & Montgomery, 2002; Menjo, 2018; Cho, 2018). This shows a common ignorance of discussing validity or reliability in composing academic articles in English intonation perception and

production. Besides, some authors also disclosed limitations related to their research designs, for instance, the small sample size, a lack of sample diversity (Hudson et al., 2022) while other authors were concerned with the quasi-longitudinal research method (Menjo, 2018), the challenge in objectively rating participants' intonation production, unbalanced testing materials (Sonsaat-Hegelheimer & Levis, 2025), settings and training effect (Cho, 2018).

In the target article pool, eight of them never discuss the directions for future research. As for the rest of them, there are so many directions for future research that it is extremely hard to categorise them according to themes and codes. Therefore, this review just summarises and reports directly by using the following words. Colantoni et al. (2016) suggested future research on the realization of English sentence types should investigate Spanish speakers' higher production of pitch accent in questions than in statements. Katsika et al. (2023) suggested future research focus on assessing the contribution of individual phonetic dimensions to production-perception relationship. Riquelme et al. (2025) advocated more research in sentence accent placement in varied linguistic contexts rather than nuclear accent. Menjo (2018) called on more perception studies while Cho (2018) urged naturalistic data to increase the task validity. Wang et al. (2010) encouraged the future study of linguistic or phonological meaning of those prominence patterns. Future research should also cover characteristics of the prosodic patterns in the different language varieties and take a pragmatic view (Swerts & Zerbian, 2010). More important directions emerged from the studies with pedagogical perspectives: to check if there is a correlation between the improvements in perception and production and if production accuracy increase with longer period of instruction (Sonsaat-Hegelheimer & Levis, 2025).

All in all, future research should concentrate on various phonetic dimensions, their functions in production-perception relationship of intonation, discrepancies of certain acoustic cues in different sentence types, diversified language varieties, pedagogical and pragmatic perspectives. Especially in the intonation training, more studies should explore different L2 speakers, contextual information, retention of intonation acquisition.

TABLE 4. Measurement validity, reliability and research design limitations

Theme	Code	Studies
Measurement validity	Validity, reliability mentioned	Sonsaat-Hegelheimer & Levis (2025); Ulbrich (2010); Stepling & Montgomery (2002); Menjo (2018); Cho (2018)
	Not mentioned	The rest of the studies
Research design limitations	Small sample size/diversity	Hudson et al. (2022); Riquelme et al. (2025); Buzan et al. (2022); Gasser et al. (2019); Menjo (2018);
	Unbalanced testing materials	Hudson et al. (2022);
	Method defects	Sonsaat-Hegelheimer & Levis (2025); Menjo (2018); Cho (2018)

DISCUSSION

The present study is a systematic literature review (SLR) of empirical studies and doctoral dissertations from inception to 2025 on the perception and production of English intonation. Following PRISMA framework, this review synthesises findings from 19 empirical studies from Web of Science (core collection), Scopus and ProQuest. In terms of research designs, most studies prefer mixed-mode research designs, which could be attributed to both the qualitative nature of human speech and the need to quantify acoustic attributes. The proficiency level of participants

normally depends on the researchers' methodological design and their research objectives, so participants of different proficiency levels were involved. Advanced speakers were chosen mainly because prosody cues belong to advanced category in L2 language learning while participants of different proficiency levels can help reveal proficiency's effect on the perception-production chain. Besides, it is statistically meaningful to sample participants of certain and representative categories. As for the participants' education background, participants from tertiary education background were widely seen, but there were no participants from elementary and participants from secondary education only appeared in one study (Colantoni et al., 2022). This indicates a common ignorance of early stages of L2 intonation acquisition. Similar instruments were adopted in that the most fundamental method to testing perception and production is through speaking and listening, which limits the instrumentation methods. However, qualitative instruments like questionnaires and surveys provide options to complement quantitative instruments. Intervention or training can be detected only in Sonsaat-Hegelheimer and Levis (2025), suggesting a rigorous research gap to apply training to English intonation perception and production.

As for research domains, Tonality is rarely mentioned while Tonicity and Tone were often studied collectively. The reason that Tonality does not seem to appeal to researchers may be attributed to the fact that Tonality, mainly indicated by pausing and lengthening as acoustic cues, varies from person to person, and there seems to be no common agreement on the test of it. The subjectivity in its nature prevents researchers from doing related studies. Besides, chunking and pauses are less meaningful in form and meaning compared with other prosodic cues. The reason for the combination of researching both, Tonicity and Tone, is that they are highly related to and dependent upon each another. Generally speaking, the nucleus marks the information focus and also signals the start of nuclear pitch movement from tonic syllable. Therefore, these two intonation features are often included in one study. For language pairing, though there was a great variety of language pairings, one study experimented on Mandarin-English speakers only, which led to a population gap for certain language groups like Mandarin EFL learners.

There have been conflicting opinions on English intonation perception and production. As for Tonicity, "no correlations or associations" (Cao & Li, 2022) contradict with "correlations" (Zhao et al., 2022) and "asymmetry" (Riquelme et al., 2025; Ali & Taha, 2022). Ortega-Llebaria & Colantoni (2014) found that L1 transfer was more evident in production than in perception, suggesting asymmetric transfer effect. In addition, asymmetry (Sonsaat-Hegelheimer & Levis, 2025) between the perception and production of English tones can be found in many studies, and the production seems to be better than perception performance for participants in many cases (Hudson et al., 2022; Puga et al., 2018; Ali & Taha, 2022; Colantoni et al., 2022). Correlational interaction may also emerge between the perception and production of specific tone types, especially for Cantonese speakers of English (Zhao et al., 2022). Noticeably, unlike Speech Learning Model (SLM) (Flege, 1995), which posits that correct sound perception leads to correct sound production, the prosody perception and production may be different and more complex depending on a variety of factors (Puga et al., 2018).

CONCLUSION

This systematic review synthesises findings from 19 empirical studies on English intonation perception and production, generating the following key themes: (1) The type of research design. Most of them would adopt mixed-mode research designs. (2) Research domain. Tonality has long been ignored, and it boasts a big potential for future research; Tonicity and Tone are the two most popular research domains, and sometimes they are intertwined in one study because of their proximity in English intonation. The correlation between perception and production, L1 transfer to English are the two major issues in English intonation studies. (3) Participants selection. Most studies sample participants from tertiary education, and their identity is solely college students. Participants from elementary or secondary education deserve more attention. (4) Intervention or teaching. Only one study in the pool discusses intervention or training. Pedagogical research should be encouraged because the meaning of doing perception-production intonation research is to ultimately apply the findings to actual learning and teaching career. (5) Measurement validity. Actually, this should be an integral part of every scientific study. Nevertheless, most studies in the pool did not mention validity or reliability tests. Therefore, to better improve our research design, especially the measurement part, it is necessary to introduce validity and reliability of the study. (6) Limitation of research designs. A larger sample size should be considered, as there seems to be a wider complaint about the small sample size. It will also be better to develop a more balanced distribution of testing materials. Every part of the testing material should account for a similar or reasonable proportion. (7) Perception-production relationship of English intonation. There have been conflicting opinions (asymmetrical or correlational) and more exploration is needed to delve into this sophisticated research area for speakers from different language backgrounds.

As for the (8) directions of future research, future research should concentrate on various phonetic dimensions, their functions in production-perception relationship of intonation, discrepancies of certain acoustic cues in different sentence types, more diversified language varieties, pedagogical and pragmatic approaches. Most importantly, in the intonation teaching field, future research should explore different L2 speakers, contextual information, retention of intonation acquisition. It is also of great significance to investigate L1 transfer or cross-linguistic influence and its effect on English intonation perception and production.

REFERENCES

- Ahmed, S. K., Mohammed, R. A., Nashwan, A. J., Ibrahim, R. H., Abdalla, A. Q., Ameen, B. M. M., & Khahir, R. M. (2025). Using thematic analysis in qualitative research. *Journal of medicine, surgery, and public health*, 6, 100198.
- Altmann, H. (2006). *The perception and production of second language stress: A cross -linguistic experimental study* [Doctoral dissertation, The University of Delaware]. ProQuest Dissertations & Theses Global. <https://www.proquest.com/dissertations-theses/perception-production-second-language-stress/docview/305322624/se-2>
- Buzan, T., Name, C., & Sosa, J. (2022). Intonational interference in English-L2 Brazilian speakers: production and perception. *Proc. SpeechProsody 2022* (pp. 828-831).
- Cao, C., & Li, A. (2022). The Realization and Perception of Narrow Focus in English Sentences by Cantonese EFL learners. *2022 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)* (pp. 560-565). IEEE.
- Cho, M. K. (2018). *An Investigation of Korean Learners' Difficulties in Using English Intonation to Express Emotion: Perception and Production* [Doctoral dissertation, University of Oxford]. ProQuest Dissertations & Theses Global. <https://www.proquest.com/dissertations-theses/investigation-korean-learners-difficulties-using/docview/2773987439/se-2>

- Colantoni, L., Klassen, G., Patience, M., Radu, M., & Tararova, O. (2022). Perception and production of sentence types by Inuktitut-English bilinguals. *Languages*, 7(3), 193.
- Colantoni, L., Klassen, G., Patience, M., Radu, M., Tararova, O. (2016). Task-effects in the L2 perception and production of English sentence types by L1 Spanish speakers. *Proc. Speech Prosody 2016*, 11-15, 10.21437/SpeechProsody.2016-3
- Cruttenden, A. (1997). *Intonation*. 2d ed. Cambridge, Cambridge Univ. Press. 10.1017/CBO9781139166973
- Demina, M. A. (2023). Selling English: Advertising as a Phonopragmatic Medium for Teaching Intonation Use. *3L: Southeast Asian Journal of English Language Studies*, 29(4), 160.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, 92(1), 233-277.
- Gasser, E., Ahn, B., Napoli, D. J., & Zhou, Z. L. (2019). Production, perception, and communicative goals of American newscaster speech. *Language in Society*, 48(2), 233-259.
- Hori, T., Akatsuka, M., & Toyama, M. (2025). Effects of observing pitch gestures on the perception of English intonation by Japanese learners of English. *Journal of Second Language Pronunciation*, 11(1), 26-45.
- Hori, T., Toyama, M., & Akatsuka, M. (2020). Perception of English intonation by Japanese learners of English. *Proceedings of the 10th International Conference on Speech Prosody 2020* (pp. 838-842).
- Hudson, T., Setter, J., & Mok, P. (2022). English intonation in storytelling: A comparison of the recognition and production of nuclear tones by British and Hong Kong English speakers. *English World-Wide*, 43(3), 357-381.
- Kang, O., Kermad, A., & Taguchi, N. (2021). The interplay of proficiency and study abroad experience on the prosody of L2 speech acts. *Journal of Second Language Pronunciation*, 7(3), 343-369.
- Kang, Y. (2024). A study on suprasegmental acquisition of Chinese EFL learners from the perspective of L1 transfer. *Addressing Global Challenges-Exploring Socio-Cultural Dynamics and Sustainable Solutions in a Changing World* (pp. 265-276). Routledge.
- Katsika, A., Jang, J., Krivokapić, J., Goldstein, L., & Saltzman, E. (2023). A hierarchy of prominence: The production and perception of focus in American English. *Proceedings of the international congress of phonetic sciences* (pp. 1696-1700).
- Levis, J. (2012). Suprasegmentals: Intonation. *The Encyclopedia of Applied Linguistics*, C.A. Chapelle (Ed.). <https://doi.org/10.1002/9781405198431.wbeal1124>
- Li, S., Wang, Y., Tang, P. T., & Feng, Y. (2024). The Role of Auditory and Visual Modality in Perception of English Statements and Echoic Question by Chinese EFL Learners. *Proc. Speech Prosody 2024* (pp. 742-746).
- Li, T., & Feng, H. (2024). Cross-linguistic Influence on Intonation Acquisition: A Study on the Production of L2 Mandarin and L3 English Intonations by Uyghur Speakers. *Proc. SpeechProsody 2024* (pp. 96-100).
- Li, Z. M., Lian, A. P., & Yodkamlue, B. (2020). Learning English intonation through exposure to resynthesized self-produced stimuli. *Gema Online Journal of Language Studies*, 20(1), 54-76.
- Liu, D., & Reed, M. (2021). Exploring the complexity of the L2 intonation system: An acoustic and eye-tracking study. *Frontiers in Communication*, 6, 627316.
- M.J. Page et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Journal of Clinical Epidemiology*. <https://doi.org/10.1016/j.jclinepi.2021.03.001>
- Menjo, S. (2018). *The Development of L2 Prosody in Japanese Speakers of English: A Quasi-Longitudinal Study*. [Doctoral dissertation, Texas A&M University-Commerce]. Dissertations & Theses Global. <https://www.proquest.com/dissertations-theses/development-l2-prosody-japanese-speakers-english/docview/2103180418/se-2>
- 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(2), 331-353.
- Puga, K., Fuchs, R., Hudson, T., Setter, J., Mok, P. (2018) The Perception-Production Link in Intonation: Evidence from German Learners of English. *Proc. Speech Prosody 2018*, 685-689, 10.21437/SpeechProsody.2018-139
- Radu, M., Colantoni, L., Klassen, G., Patience, M., Pérez-Leroux, A. T., & Tararova, O. (2020). The perception and interpretation of sentence types by L1 Spanish–L2 English speakers. *Linguistic Approaches to Bilingualism*, 10(4), 499-529.
- Riquelme, C., Pereira-Reyes, Y., & Véliz-Campos, M. (2025). Unpacking the Possible Effects of Working Memory and English Proficiency on Nuclear Accent Perception and Production in Chilean Learners of English. *Language Teaching Research Quarterly*, 49, 41-58.

- Sonsaat-Hegelheimer, S., & Levis, J. (2025). Is intonation learnable in the classroom? Evidence from Turkish learners of English. *Language Teaching Research*, 13621688241312502.
- Steppling, M. L., & Montgomery, A. A. (2002). Perception and production of rise-fall intonation in American English. *Perception & psychophysics*, 64(3), 451-461.
- Sundus Muhsin Ali & Mariam Saad Taha (2022). Iraqi EFL College Students' Recognition and Production of Intonation. *Res Militaris*, 12 (2) 6085-6092. <https://resmilitaris.net/issue-content/iraqi-e-fl-college-students-recognition-and-production-of-intonation-529>
- Swerts, M., & Zerbian, S. (2010). Intonational differences between L1 and L2 English in South Africa. *Phonetica*, 67(3), 127-146.
- Tian, Q., Yap, N. T., & Ng, C. F. (2024). The Effect of English Proficiency on the Production of English Intonation by Chinese EFL Learners. *Pertanika Journal of Social Sciences & Humanities*, 32(3).
- Ulbrich, C. (2010). Belfast Intonation in L2 speech. *Proceedings of Speech Prosody* (Vol. 2010).
- Vaissière, J. (2005). Perception of intonation. *The handbook of speech perception*, 236-263.
- Van, P. T. T., Lien, L. K., & Khang, N. D. (2024). Effects of Video-Based Shadowing on Suprasegmental Features: EFL Learners' Pronunciation Performance and Attitudes. *rEFLections*, 31(3), 896-924.
- Wang, X., Li, A., & Ji, X. (2010). Perception and production of prominence distribution patterns of Chinese EFL Learners. *Proc. SpeechProsody 2010* (pp. paper-046).
- Wells, J.C. (2006). *English Intonation: An Introduction*. Cambridge University Press.
- Zhao, C., Xiong, Z., & Li, A. (2020). Using multimodal methods in L2 intonation teaching for Chinese EFL learners. *2020 23rd Conference of the Oriental COCOSDA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSDA)* (pp. 100-105). IEEE.
- Zhao, C., Xiong, Z., & Li, A. (2022). Production and Perception of Intonation Features by Cantonese EFL Learners. *2022 25th Conference of the Oriental COCOSDA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSDA)* (pp. 1-5). IEEE.
- Zhi, N., Li, A., & Zhao, C. (2024). The effect of visual displays on the mastery of the forms and functions of English intonation by Chinese L2 students. *Lingua*, 311, 103825.