Language Teaching (2025), 1–18
doi:10.1017/S0261444825101006

CAMBRIDGE UNIVERSITY PRESS

RESEARCH TIMELINE

L2 speech intelligibility

Okim Kang[1], Yuna Bae[1] and Yongzhi Miao[2]

[1]English Department, Northern Arizona University, Flagstaff, AZ, USA and [2]Department of Classical and Modern Languages and Literatures, Texas Tech University, Lubbock, TX, USA
Corresponding author: Okim Kang; Email: okim.kang@nau.edu

(Received: 15 August 2024; accepted: 2 October 2025)
Keywords: L2 pronunciation; L2 speech; second language intelligibility

1. Introduction

One may intuitively think of second language (L2) intelligibility as how well listeners understand a speaker using an L2. It may be commonly perceived as relating to the understanding and comprehension of speech. Then, this conception of L2 intelligibility has been increasingly popular within the domain of L2 speech as well as language teaching, learning, and assessment. Certainly, the publication of Munro and Derwing's (2011) research timeline on the topic of accent and intelligibility in pronunciation research in *Language Teaching* has been an important addition to this movement. While we have seen challenges stemming from the limited examinations and the broad yet nebulous scope of applied phonetics and phonology (Munro & Derwing, 2011), the topic of L2 intelligibility is now undoubtedly recognized as a key topic in the field. In fact, intelligibility is critical for effective communication, which is manifested not only in the context of language classrooms but also in everyday L2 usage. Therefore, it has naturally moved beyond a sole focus on speech signals (Koffi, 2021) and has emerged as both an instructional focus (Levis, 2005) and an assessment criterion (Kang et al., 2023). It has also become central to discussion about the globalization of L2 English, particularly in reference to World Englishes (WE) and English as a Lingua Franca (ELF) (Jenkins et al., 2011).

However, although a substantial body of literature has attempted to address the concept of L2 intelligibility, the lack of a unitary definition seems to persist. Intelligibility has been defined and measured in many different ways (Jenkins, 2002*; Kang et al., 2018*; Munro, 2008), suggesting that intelligibility is a complex construct affected by multiple dimensions, including perceptual, linguistic, and acoustic features of speech. In 1985, Smith and Nelson synthesized 163 research studies on comprehension and intelligibility published between 1950 and 1985. They divided the broad concept of intelligibility (i.e., international intelligibility) into three sub-categories: intelligibility, comprehensibility, and interpretability. They suggested defining intelligibility more narrowly as the recognition of words or utterances. This distinction was made to address the various interpretations of intelligibility, requiring clearer terminology for more precise analysis and examination.

Forty years later, a need remains to establish a clearer definition of L2 intelligibility, as what constitutes intelligibility varies from context to context. The very features that comprise the intelligibility of specific accented varieties have not yet been fully established (Kang et al., 2020b*). Also, intelligibility measures have been debated from various disciplines, including speech pathology (Liss, 2024; Xue et al., 2024) and acoustic phonetics (Babel & Russell, 2015; Pérez-Ramón et al., 2022; Winters &

*Indicates that the complete reference is listed in the following timeline.

O'Brien, 2013*). Therefore, discussion on speech intelligibility is needed to better understand such a multifaceted construct both conceptually and methodologically.

Following Munro and Derwing's (1995*) and Derwing and Munro's (1997*) seminal work, three speech constructs have come to dominate the L2 pronunciation literature: intelligibility, comprehensibility, and accentedness. Intelligibility is referred to as the degree to which a listener actually understands a speaker's intended message, accentedness as phonological differences between L2 and target language, and comprehensibility as how easy or difficult a speaker is to understand, focusing on processing efforts. In other words, intelligibility can be defined as the accuracy to which a listener can identify the intended utterances produced by a speaker. This distinction has led to a wide acceptance that these constructs are at least partially independent (Derwing & Munro, 1997). That is, an L2 speaker's speech can be heavily accented but still be largely intelligible, although some research suggests that intelligibility usually decreases as the degree of accent increases (Rogers et al., 2004). On the other hand, a lack of intelligibility does not necessarily indicate a weak accent. In fact, speakers who are easily understood may still be perceived as having a strong accent (Kang et al., 2018*).

Another important characteristic of intelligibility is that it serves a vital communicative function in our everyday interactions, and the success of information transfer depends on the mutual intelligibility between the speaker and the listener. Understanding is an interactive process that requires engagement from both speakers and listeners, rather than relying solely on one party (Smith & Nelson, 2019). Given that communication is a collaborative activity requiring active participation from both speakers and listeners, sharing responsibility for a successful interaction is key to global communication (Clark & Wilkes-Gibbs, 1986; Derwing et al., 2014*). Put differently, a speaker's intelligibility can depend on situational, social, and cultural factors (Smith & Nelson, 2019).

Much research has also examined listeners' language and linguistic backgrounds that affect L2 intelligibility. Gass and Varonis (1984*) demonstrated listeners' language experiences (e.g., familiarity with the topic, accent, speaker) were strongly correlated with their judgments of L2 intelligibility. Note that while the authors used the term comprehensibility, their measure was more in line with intelligibility measures. One factor that also contributes to greater tolerance of listeners is known as "the interlanguage speech intelligibility benefit (ISIB)" (Bent & Bradlow, 2003). This asserts that L2 listeners may have benefits in interpreting specific acoustic-phonetic features of an L2 that are matched with their own L1s. Although some studies showed mixed results (Major et al., 2002) or minimal effects of listeners' L1 (Munro et al., 2006*), other recent studies have demonstrated consistent shared-L1 effects among L2 listeners with certain L1 backgrounds, i.e., especially among listeners from outer circle countries (e.g., India or South Africa) where English is spoken as an official and second language (e.g., Kang et al., 2019, 2023; Shin et al., 2021). To explain this, some scholars have argued that shared L1 may not directly influence understanding but may do so in more complex and indirect ways (Miao, 2023).

Currently, research on L2 intelligibility seems to be at a turning point, expanding its scope and depth of understanding. Previously, understanding diverse accented speech was a central concern in the globalized world; thus, most research primarily investigated this construct based on human judgments. However, with the rise of interaction with artificial intelligence (AI) chatbots (Labadze et al., 2023), learners or speakers are now at the stage of prioritizing intelligible speech even in their communications with AI (Moussalli & Cardoso, 2020*). Furthermore, the evolution of automatic speech recognition (ASR) capabilities in AI technology has spurred an increase in research focused on ASR-based L2 intelligibility (Inceoglu et al., 2023*). This means that we are entering the initial stage of a new era in speech intelligibility. Accordingly, the current timeline on L2 intelligibility is timely and appropriate at both practical and pedagogical levels. The field stands to benefit from carefully compiled and scrutinized studies, along with their key findings.

Note that the current research timeline stands out distinctly from its predecessor. While Munro & Derwing (2011) examined studies concerning classroom pronunciation instruction, with a focus on

https://doi.org/10.1017/S0261444825101006 Published online by Cambridge University Press

both accent and intelligibility, the objective of our current timeline is to center the research solely on L2 intelligibility itself. Our aim is to delve into definitions, measures, and constructs of L2 intelligibility which have rapidly evolved over the last four decades. Therefore, our timeline can shed light on the trajectory that L2 intelligibility has followed up to 2024. We further aim to provide readers with insights into potential future directions for L2 intelligibility research and practices. We encourage readers to reflect on how the field has interpreted the notion of intelligibility and its implementation in practice.

2. Coding process

Our initial examination relied on *Google Scholar* and library database searches as the primary sources for pertinent studies. Throughout this search process, key terms such as "L2 intelligibility" and "speech intelligibility" were employed. It is important to note that our initial search was restricted to publications in academic journals, conference proceedings, and book chapters. Theses and dissertations were excluded. Moreover, only documents published in English were investigated due to length constraints and practicality considerations. As a result, a total of 117 studies were primarily identified. Subsequently, specific criteria were applied to refine and finalize the selection of the studies.

To begin with, our timeline exclusively incorporated primarily empirical investigations that had been published in peer-reviewed journals, meaning that conceptual or argumentative manuscripts were excluded. Additionally, a decision was made that conference proceedings and book chapters be omitted from this process to ensure that all included studies underwent rigorous peer review. In terms of the selection of journals, we confined our choice to those indexed in the Social Sciences Citation Index (SSCI) by Clarivate Analytics (https://mjl.clarivate.com/home). Consequently, articles from non-indexed journals focusing on L2 intelligibility were omitted. However, there were three exceptions to this criterion: the journals *Speech Communication, Journal of the Acoustical Society of America*, and *Journal of Second Language Pronunciation* were included despite the first two belonging to the Science Citation Index Expanded (SCIE) and the last one to the Emerging Sources Citation Index (ESCI), respectively. This choice was made because SCIE is considered to correspond to SSCI, and the latter journal holds significance in the field of L2 speech in spite of its indexing status. In the case of review papers, a few were retained under two conditions: (1) they must contain empirical research, and (2) must have been published in peer-reviewed journals. Conversely, non-empirical review papers, particularly those that mainly discuss theoretical aspects, were removed. Moreover, our selections were limited to papers that encompassed pedagogical implications due to their relevance to the journal of *Language Teaching's* aim and scope. As a result, we have narrowed down our timeline to 50 studies.

3. The current research timeline

During the coding process, we noticed an increase in interest in L2 intelligibility in recent years. The current research timeline includes a total of 50 studies, with 16 conducted before 2011 and 34 after 2012. Surprisingly, 21 of these studies have been conducted since 2020 alone. This trend highlights the growing attention to L2 speech intelligibility within the field of Applied Linguistics.

Drawing from the discussions outlined above, we identified the major themes of our research timeline. These themes are categorized into three main areas based on their focus: *what* (i.e., the research focus), *which* (i.e., the perspectives from which intelligibility is approached), and *how* (i.e., the way intelligibility is defined and operationalized). The first theme pertains to research focus (**F**), which signifies *what* the studies have investigated. This theme is particularly pertinent to research questions. Within this theme, there are three sub-themes. One of these sub-themes revolves around the relationship between intelligibility and other speech constructs such as comprehensibility and accentedness (**F-SpC**). While earlier studies predominantly focused on the interplay among speech constructs (e.g.,

https://doi.org/10.1017/S0261444825101006 Published online by Cambridge University Press

Derwing & Munro, 1997*; Munro & Derwing, 1995*), more recent studies have addressed comprehensibility and accentedness alongside other aspects of intelligibility such as intelligibility-related phonetic features or contextual factors (e.g., Ali, 2023*; Huensch & Nagle, 2021*, 2023*; Jułkowska & Cebrian, 2015*; Kennedy & Trofimovich, 2008*, Matsuura et al., 1999*). Studies falling under the second sub-theme examined the variables that affect intelligibility (**F-Var**). These variables may include factors related to the listener (e.g., Field, 2005*; Kang et al., 2020a*; Kennedy & Trofimovich, 2008*; Matsuura, 2007*; Munro et al., 2006*), the speaker (e.g., Field, 2005*; Hahn, 2004*; Munro et al., 2006*; Setter, 2006*; Zielinski, 2008*), or specific speech features (e.g., Deterding & Kirkpatrick, 2006*; Emara & Shaker, 2024*; Kang et al., 2020b*). It is noteworthy that the majority of timeline studies encompass this sub-theme, indicating a strong desire within the field to examine intelligibility and explain the characteristics of intelligible speech. The last sub-theme is directly linked with L2 intelligibility itself (**F-Int**). This sub-theme includes studies whose primary objective is to define, refine, or explore L2 intelligibility as a construct. Such studies may involve conceptual discussions of intelligibility (e.g., Smith & Rafiqzad, 1979*), empirical efforts to establish thresholds that distinguish intelligible from unintelligible speech (e.g., Kang et al., 2020b*), or comparisons of various methods used to measure intelligibility (e.g., Brodkey, 1972*; Kang et al., 2018*).

The second major theme, "which," refers to the various perspectives (**P**) from *which* the concept of intelligibility is approached. Given that intelligibility is explored across diverse disciplines and viewpoints, this theme provides a comprehensive framework for understanding its conceptual scope. In this timeline, intelligibility can be interpreted from six different angles as follows: (1) an instructional perspective (**P-Inst**) (e.g., Yenkimaleki & Van Heuven, 2022*), where intelligibility is viewed as an essence of the intelligibility principle, which prioritizes effective communication and listener understanding. This contrasts with the nativeness principle, which emphasizes native-like pronunciation as the ultimate goal (Levis, 2005). In addition, intelligibility can be explicated in relation to (2) technology (**P-Tech**), (3) World Englishes (**P-WEng**), and (4) social perception (**P-SocP**). To be specific, technology such as ASR is often utilized to assess intelligibility (e.g., Emara & Shaker, 2024*; Inceoglu et al., 2023*; Mroz, 2018*, 2020*), while it can also be construed in terms of international intelligibility (Sewell, 2013*), and listeners' negative perceptions of L2 intelligible speech (e.g., Hendriks et al., 2023*; Lee & Bailey, 2023*; Rubin, 1992). As a concept related to speech, intelligibility can also be analyzed from either (5) a phonetics perspective (**P-APhon**) – encompassing acoustic, articulatory, and auditory phonetics – or (6) an L2 pronunciation perspective (**P-Pron**). Although these perspectives may appear similar, they differ in focus and methodology. The former views intelligibility as a physical phenomenon, analyzing segmental features (e.g., vowels, consonants) and measurable speech properties using acoustic and articulatory analysis, including parameters like F0 (e.g., Aoyama et al., 2023*; Jin & Liu, 2014*; Kawase et al., 2014*; Winters & O'Brien, 2013*). Conversely, the latter defines intelligibility as a perceptual construct, prioritizing listener-based assessments over acoustic measurements. While it also considers pronunciation features, it focuses more on how speech is interpreted and judged by listeners beyond purely phonetic dimensions (Yates, 2017). Many applied linguistics studies (e.g., Field, 2005*; Hahn, 2004*; Kennedy & Trofimovich, 2008*) fall into this category.

Our third major theme focuses on the operationalization of L2 speech intelligibility (**O**), exploring *how* it is defined and implemented in each study. As previously mentioned, there is no singular, universally accepted definition of intelligibility. Consequently, researchers in the field have employed various measurement methods, reflecting the absence of standardized and widely accepted approaches. Considering that the operationalization of L2 intelligibility is based on its definition, we identified the final major theme as examining how scholars operationalize intelligibility. In other words, regardless of the research focus (F) or perspective (P), if intelligibility is operationalized in a specific way, the corresponding O theme should also be applied. For example, a study that investigates intelligibility thresholds (F-Int) and uses transcriptions as the measurement method should be coded as both F-Int and O-Tran. In many studies in our timeline, L2 intelligibility was operationalized by using a transcription method (**O-Tran**). This approach involves measuring intelligibility based on

https://doi.org/10.1017/S0261444825101006 Published online by Cambridge University Press

Table 1. Overview of the timeline studies

Themes	Number of studies	Percentage of studies (%)
*Research focus (F)		
F-SpC	9	18
F-Var	39	78
F-Int	11	22
*Perspectives (P)		
P-Inst	7	14
P-Tech	7	14
P-WEng	10	20
P-SocP	1	2
P-APhon	5	10
P-Pron	32	64
*Operationalization (O)		
O-Tran	33	66
O-Scal	5	10
O-Othr	14	28
Target language (TL)		
TL-Engl	41	82
TL-Othr	9	18

*Totals may not add up to 50, as some studies employed more than one method. Indicates that the complete reference is listed in the following timeline.

the number of words or sentences listeners correctly understand. Alternatively, intelligibility could be operationalized through listeners' scalar ratings (**O-Scal**). In this case, intelligibility is assessed by listeners' subjective judgments of their own understanding of L2 speech (e.g., Lascotte & Tarone, 2022*; Moussalli & Cardoso, 2020*; Murphy, 2014*; Saito & Van Poeteren, 2012*; Yenkimaleki & Van Heuven, 2021*). If the studies employed alternative intelligibility measures (e.g., Aoyama et al., 2023*; Deterding & Kirkpatrick, 2006*; Jin & Liu, 2014*; Kawase et al., 2014*) or incorporated a combination of different methods (e.g., Emara & Shaker, 2024*; Moussalli & Cardoso, 2020*), we coded them as **O-Othr**.

In sum, the themes in our timeline have been coded as follows:

F: Research focus (*what* the studies investigated)
- **F-SpC**: relationship with other speech constructs (i.e., comprehensibility, accentedness)
- **F-Var**: variables/features affecting intelligibility
- **F-Int**: intelligibility itself (e.g., measures, thresholds)

P: various perspectives from *which* the concept of intelligibility is approached
- **P-Inst**: instructional approach (intelligibility principle vs. nativeness principle)
- **P-Tech**: technology-related (ASR)
- **P-WEng**: World Englishes
- **P-SocP**: social perception (social discrimination, bias)
- **P-APhon**: acoustic phonetics
- **P-Pron**: pronunciation

O: operationalization of L2 intelligibility (*how* it is defined and operationalized)
- **O-Tran**: transcription methods (includes word/sentence transcription)
- **O-Scal**: scalar rating
- **O-Othr**: other methods

https://doi.org/10.1017/S0261444825101006 Published online by Cambridge University Press

In addition to the primary themes, our timeline encompassed an additional supplementary category. Given that our timeline focuses on L2 speech, each study naturally had a designated target language for intelligible speech. Therefore, we divided the themes into **TL-Engl** (English as the target L2) and **TL-Othr** (other languages as the target L2).

TL: target language

TL-Engl: English is a target L2

TL-Othr: other languages are the target L2

In terms of the coding process, the second author, serving as the primary coder, coded all themes, while the third author independently coded 12.24% of the timeline studies. Interrater agreement was calculated as the proportion of shared codes to total codes across the double-coded sample (29/30). A 97% interrater agreement was achieved, ensuring reliability in the coding process. When disagreements occurred, the authors met and discussed to reach a consensus. Table 1 presents an overview of the timeline studies.

The present timeline aims to comprehensively incorporate prominent contributions to L2 speech intelligibility, providing our readers with an overview of its change and offering deeper insight into how perspectives, definitions, and measures of L2 intelligibility have evolved. Therefore, it should be noted that our annotations in the timeline specifically set out to focus on L2 intelligibility and its operationalization.

References

Babel, M., & Russell, J. (2015). Expectations and speech intelligibility. *The Journal of the Acoustical Society of America*, *137*(5), 2823–2833.

Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America*, *114*(3), 1600–1610.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*(1), 1–39.

Hirschi, K., & Kang, O. (2023). How many eaters can be enough: G Theory applied to assessment and measurement of L2 speech perception. *Language Teaching Research Quarterly*, *37*, 213–230. https://doi.org/10.32038/ltrq.2023.37.12

Isaacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility: Identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition*, *34*(3), 475–505.

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

Jenkins, J., Cogo, A., & Dewey, M. (2011). Review of developments in research into English as a lingua franca. *Language Teaching*, *44*(3), 281–315.

Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *Modern Language Journal*, *94*(4), 554–566.

Kang, O., & Rubin, D. L. (2009). Reverse linguistic stereotyping: Measuring the effect of listener expectations on speech evaluation. *Journal of Language and Social Psychology*, *28*(4), 441–456.

Kang, O., Thomson, R., & Moran, M. (2019). The effects of international accents and shared first language on listening comprehension tests. *TESOL Quarterly*, *53*(1), 56–81.

Kang, O., Yan, X., Kostromitina, M., Thomson, R., & Isaacs, T. (2023). Fairness of using different English accents: The effect of shared L1s in listening tasks of the Duolingo English test. *Language Testing*, 02655322231179134. https://doi.org/10.1177/02655322231179134

Koffi, E. (2021). *Relevant acoustic phonetics of l2 english: Focus on intelligibility* (1st ed.). CRC Press. https://doi.org/10.1201/9781003106418

Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: Systematic literature review. *International Journal of Educational Technology in Higher Education*, *20*(1), 56.

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, *39*(3), 369.

Liss, J. (2024). Speech intelligibility. In M. J. Ball, N. Müller, & E. Spencer (Eds), *The handbook of clinical linguistics* (2nd ed., pp. 605–614). Wiley-Blackwell.

Major, R. C., Fitzmaurice, S. F., Bunta, F., & Balasubramanian, C. (2002). The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly*, *36*(2), 173–190.

Miao, Y. (2023). The relationship among accent familiarity, shared L1, and comprehensibility: A path analysis perspective. *Language Testing*, *40*(3), 723–747.

Munro, M. J. (2008). Foreign accent and speech intelligibility. In J. G. H. Edwards & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 193–218). John Benjamins.

https://doi.org/10.1017/S0261444825101006 Published online by Cambridge University Press

Munro, M. J., & Derwing, T. M. (2011). The foundations of accent and intelligibility in pronunciation research. *Language Teaching*, *44*(3), 316–327.

Pérez-Ramón, R., García Lecumberri, M. L., & Cooke, M. (2022). Foreign accent strength and intelligibility at the segmental level. *Speech Communication*, *137*, 70–76.

Rogers, C. L., Dalby, J., & Nishi, K. (2004). Effects of noise and proficiency on intelligibility of Chinese-accented English. *Language and Speech*, *47*(2), 139–154.

Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education*, *33*(4), 511–531.

Shin, S. Y., Lee, S., & Lidster, R. (2021). Examining the effects of different English speech varieties on an L2 academic listening comprehension test at the item level. *Language Testing*, *38*(4), 580–601.

Smith, L. E., & Nelson, C. L. (2019). World Englishes and issues of intelligibility. In C. L. Nelson, Z. G. Proshina & D. R. Davis (Eds.), *The handbook of world Englishes* (1st ed, pp. 430–446). Wiley. https://doi.org/10.1002/9781119147282.ch24

Xue, W., Van Hout, R., Cucchiarini, C., & Strik, H. (2023). Assessing speech intelligibility of pathological speech: Test types, ratings and transcription measures. *Clinical Linguistics & Phonetics*, *37*(1), 52–76.

Yates, L. (2017). Learning how to speak: Pronunciation, pragmatics, and practicalities in the classroom and beyond. *Language Teaching*, *50*(2), 227–246.

https://doi.org/10.1017/S0261444825101006 Published online by Cambridge University Press

Year	References	Annotations	Theme
1972	Brodkey, D. (1972). Dictation as a measure of mutual intelligibility: A pilot study. *Language Learning, 22*(2), 203–217.	**Brodkey** proposed dictation as a useful tool for measuring mutual intelligibility across native and foreign accents, providing foundational support for transcription as a commonly used intelligibility measure in the field (e.g., Munro & Derwing, 1995[1]; Munro et al., 2006). In this study, intelligibility was conceptualized as the extent to which individuals or groups can effectively understand and be understood by speakers of different languages or dialects, and it was operationalized through dictation results. However, the study did not focus solely on individual word-for-word errors. Instead, dictation accuracy was determined by considering the presence of thought-groups within each dictation line. The author particularly recommended dictation as an effective methodology for TESL (Teaching English as a Second Language) teacher-trainees.	F-Var F-Int P-Pron O-Tran TL-Engl
1979	Smith, L. E., & Rafiqzad, K. (1979). English for cross-cultural communication: The question of intelligibility. *TESOL Quarterly, 13*(3), 371.	**Smith and Rafiqzad** sought to explore the degree of intelligibility across native and nonnative accents of educated English using a large sample size (n=1,368). They defined intelligibility as the capacity to understand words within the context of a naturally spoken or read sentences. Hence, they conducted a Cloze test, where listeners filled in blanks with what they heard. The results reported consistent rankings in the degree of intelligibility among various accented speeches and further revealed that native speakers' speech is not always the easiest for listeners to understand, suggesting that listeners can perceive educated English with diverse accents as intelligible. Interestingly, these findings are corroborated by later work exploring the threshold of intelligibility (Kang et al., 2020b).	F-Int P-Pron O-Tran TL-Engl
1984	Gass, S., & Varonis, E. M. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning*, 34(1), 65–87.	This pioneering study set out to examine the impact of familiarity factors on the intelligibility of nonnative speech to native listeners. Although the term "comprehensibility" was used, the study predominantly focused on L2 intelligibility. **Gass and Varonis** explicitly stated their interest in assessing listeners' understanding of L2 speech, employing transcription methods rather than relying on subjective judgments of comprehension ease. This study provided clear evidence that familiarity variables significantly influenced the intelligibility of L2 speech for native listeners, while also laying the groundwork for future studies exploring the relationship between listeners' familiarity and their judgments of speech constructs (e.g., Derwing & Munro, 1997; Matsuura et al., 1999).	F-Var P-Pron O-Tran TL-Engl
1995	Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*(1), 73–97.	This seminal work empirically demonstrated the distinction of the three constructs: accentedness, comprehensibility, and intelligibility. **Munro and Derwing**'s argument that these are related but partially independent constructs provided a foundation for many pronunciation research and teaching practices.	F-SpC P-Pron O-Tran TL-Engl
1997	Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition, 19*(1), 1–16.	**Derwing and Munro** extended their precursor work (Munro & Derwing, 1995) to include English accents from speakers of four L1 backgrounds: Cantonese, Japanese, Polish, and Spanish. This study also took into account speaker proficiency and listener familiarity. Results supported the previous study that regardless of speaker proficiency, accentedness, comprehensibility, and intelligibility still emerged as independent constructs, although the weight of linguistic features contributing to these constructs differed. Moreover, the study found that accent identification ability was influenced by listeners' accent familiarity.	F-SpC F-Var P-Pron O-Tran TL-Engl

(*Continued*)

(*Continued.*)

Year	References	Annotations	Theme
1999	Matsuura, H., Chiba, R., & Fujieda, M. (1999). Intelligibility and comprehensibility of American and Irish Englishes in Japan. *World Englishes, 18*(1), 49–62.	In their study, **Matsuura et al.** explored how accent familiarity influenced listeners' judgments. Unlike Gass and Varonis (1984), who focused on native listeners' accent familiarity, this study examined nonnative listeners' judgments on the intelligibility and comprehensibility of both familiar and unfamiliar English accents. Intelligibility was assessed through a transcription task, specifically a cloze dictation. The results revealed that intelligibility and comprehensibility did not necessarily correlate with each other, consistent with previous findings (see Derwing and Munro, 1997; Munro and Derwing, 1995). Additionally, it was found that familiarity was a significant variable affecting listeners' comprehensibility ratings but not their dictation scores (i.e., intelligibility). The study's findings indicate that L2 speech can be intelligible even to nonnative listeners who are not familiar with the speaker's accent.	F-SpC F-Var P-WEng O-Tran TL-Engl
2000	Derwing, T. M., Munro, M. J., & Carbonaro, M. (2000). Does popular speech recognition software work with ESL speech? *TESOL Quarterly, 34*(3), 592.	**Derwing et al.** sought to explore the potential of Automatic Speech Recognition (ASR) software in identifying pronunciation challenges among L2 speakers. Their study involved assessing the software's transcription accuracy and comparing it with evaluations by human listeners on intelligibility, comprehensibility, and accentedness. The transcriptions generated by both ASR and human listeners for both native and nonnative speech were compared. The findings revealed that ASR transcriptions, particularly in nonnative speech, were less accurate compared to human listeners. Additionally, the study found no significant correlation between ASR transcription accuracy and listeners' intelligibility scores. The authors advocated for cautious utilization of ASR in L2 classrooms.	F-Int P-Tech O-Tran TL-Engl
2002	Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. *Applied Linguistics, 23*(1), 83–103.	Jenkins underscores the growing role of English as an International Language (EIL), where nonnative speakers (NNS) interact more frequently with other NNSs than with native speakers (NSs). Intelligibility was measured by analyzing listeners' comprehension of NNS speech and instances of communication breakdown, aligning with the approaches taken by Deterding & Kirkpatrick (2006) and Sewell (2013). Drawing from three datasets, she examined phonological features that either facilitated or impeded intelligibility. The findings revealed that pronunciation errors (e.g., consonant substitutions, misplaced tonic stress, and vowel length errors) were the primary causes of intelligibility breakdowns. This study is particularly significant as it led to the development of the Lingua Franca Core (LFC), a practical and effective pronunciation syllabus designed to enhance intelligibility in EIL contexts.	F-Var P-WEng P-Pron O-Othr TL-EN
2004	Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly, 38*(2), 201–223.	This research has made an important contribution to the field as it sought to empirically demonstrate the relationship between sentence stress and intelligibility. Specifically, **Hahn** investigated the effect of sentence stress placement on intelligibility and discovered the importance of correct sentence stress. Findings provided implications for training for international teaching assistants and pedagogical implications in L2 language classrooms.	F-Var P-Pron O-Othr TL-Engl
2005	Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly, 39*(3), 399–423.	While Hahn (2004) highlighted the importance of sentence stress, **Field** investigated the effect of lexical stress on intelligibility. Lexical stress in this study was operationalized by the change of vowel quality and the stress location shift to another syllable. Findings suggested that intelligibility was reduced when lexical stress was not correctly realized. This finding held true for both L1 as well as L2 English listeners. The study provided practical implications for L2 pedagogy and acknowledged both L1 and L2 English users who actively take part in global communication.	F-Var P-Pron O-Tran TL-Engl

(*Continued*)

(*Continued.*)

Year	References	Annotations	Theme
2006	Deterding, D., & Kirkpatrick, A. (2006). Emerging South-East Asian Englishes and intelligibility. *World Englishes, 25*(3–4), 391–409.	The purpose of **Deterding and Kirkpatrick**'s study was to investigate the linguistic features of the ASEAN countries and their effect on intelligibility amongst the speakers in spontaneous conversations in small groups. In this study, intelligible speech was defined as smooth communication with minimal breakdowns and was measured based on the flow and interruptions in conversation, which differed significantly from the conventional transcription-based method. Results suggested that shared pronunciation features such as reduced initial aspiration, stressed pronouns, and heavy end-stress did not hinder intelligibility, while non-shared features (e.g., [ɑ:], /r/, [n], [ʃ], [t]) could impede intelligibility. This study is important because it explored intelligibility in ASEAN English as a lingua franca context.	F-Var P-WEng P-Pron O-Othr TL-Engl
2006	Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition, 28*(1), 111–131.	Expanding the research by Derwing and Munro (1997), **Munro et al.** investigated the effect of listener L1 background on the intelligibility of speakers of different L1 background. Findings largely supported that listener L1 background did not contribute too much to intelligibility and that properties of the L2 speech itself were the main contributors. This study acknowledged that both speaker-based and listener-based variables could contribute to speech perception.	F-Var P-WEng P-Pron O-Tran TL-Engl
2006	Setter, J. (2006). Speech rhythm in world Englishes: The case of Hong Kong. *TESOL Quarterly, 40*(4), 763.	As previous studies have shown that certain suprasegmental features, such as lexical stress (Field, 2005) and sentence stress (Hahn, 2004), influence the intelligibility of L2 speech, **Setter** focused on syllable duration as a measure of speech rhythm in Cantonese-accented English in Hong Kong. This article suggested some similarities of syllable duration across speakers and differences across speakers in terms of the relative duration of tonic, stressed, unstressed, and reduced syllables, which can cause intelligibility problems. The results were discussed in terms of L1 transfer, impact on intelligibility, and pedagogical implications.	F-Var P-WEng P-Pron O-Othr TL-Engl
2007	Matsuura, H. (2007). Intelligibility and individual learner differences in the EIL context. *System, 35*(3), 293–304.	Expanding upon Matsuura et al. (1999), **Matsuura** examined the intelligibility and comprehensibility of American English and HK English to Japanese EFL students in relation to listener individual differences such as familiarity (Gass & Varonis, 1984), anxiety, and perceived competence. The study was interesting in that it explored L2 English speakers' perception of both L1 and L2 speech. Findings supported that all the listener individual differences mentioned above were correlated with their intelligibility scores.	F-Var P-WEng O-Tran TL-Engl
2008	Kennedy, S., & Trofimovich, P. (2008). Intelligibility, comprehensibility, and accentedness of L2 Speech: The role of listener experience and semantic context. *Canadian Modern Language Review, 64*(3), 459–489.	This research broadens the work of Gass and Varonis (1984) by exploring the influence of listener experience (familiarity) and semantic context on intelligibility of L2 speech. The findings of **Kennedy and Trofimovich** found that experienced listeners with greater exposure to L2 speech transcribed more accurately compared to those with little exposure. Additionally, the availability of semantic context improved the intelligibility of L2 speech. Overall, the results support that intelligibility depends on both the speech proprieties as well as listener idiosyncrasies.	F-Var P-Pron O-Tran TL-Engl

(*Continued*)

(*Continued.*)

Year	References	Annotations	Theme
2008	Zielinski, B. W. (2008). The listener: No longer the silent partner in reduced intelligibility. *System, 36*(1), 69–84.	**Zielinski** focused on listeners from a rather qualitative approach and examined how three L1 English listeners transcribed L2-accented English speech. Observations indicated that listeners consistently and heavily rely on speakers' production of segments in strong syllables and syllable stress patterns, adding depth to the findings of Field (2005). This study adopted a qualitative perspective in investigating intelligibility and helped enhance language pedagogy in L2 listening and speaking by highlighting the important role of listener.	F-Var P-Pron O-Tran O-Othr TL-Engl
2012	Hayes-Harb, R., & Watzinger-Tharp, J. (2012). Accent, intelligibility, and the role of the listener: Perceptions of English-accented German by native German speakers. *Foreign Language Annals, 45*(2), 260–282.	In line with Gass & Varonis (1984) and Kennedy & Trofimovich (2008), **Hayes-Harb and Watzinger-Tharp** sought to shed light on the relationship between intelligibility and accentedness by investigating how listener background factors influenced their judgments of intelligibility and accentedness. They utilized an auditory word identification task to operationalize intelligibility, tasking native listeners with identifying words from audio recordings of L2 speakers. While most previous studies have primarily focused on L2 English, this study specifically investigated L2 German. The results echoed previous findings, indicating that an L2 accent does not necessarily reduce intelligibility (Derwing & Munro, 1997; Munro and Derwing, 1995).	F-SpC F-Var P-Pron O-Othr TL-Othr
2012	Saito, K., & Van Poeteren, K. (2012). Pronunciation-specific adjustment strategies for intelligibility in L2 teacher talk: Results and implications of a questionnaire study. *Language Awareness, 21*(4), 369–385.	This is a questionnaire-based study conducted by **Saito and Van Poeteren** to investigate teachers' reported classroom behavior, especially focusing on pronunciation-specific adjustments aimed at enhancing mutual intelligibility in EFL classrooms. To be specific, the questionnaire (a) investigated which pronunciation features teachers prioritize for their students' intelligible speech and (b) identified the adjustment strategies they employ to enhance intelligibility in classroom input. Categories and frequencies of various strategies were identified, offering implications for L2 classroom practice.	F-Var P-Inst P-Pron O-Scal TL-Engl
2013	Sewell, A. (2013). Language testing and international intelligibility: A Hong Kong case study. *Language Assessment Quarterly, 10*(4), 423–443.	**Sewell** sought to offer implications for language testing by comparing test examiners' comments on L2 speakers' responses to speaking tasks with criteria for international intelligibility (Jenkins, 2000). Jenkins's (2000) Lingua Franca Core (LFC) was used as the criterion for intelligibility, and, similar to Deterding & Kirkpatrick (2006), phonological features were investigated. The results suggested that examiner notes on segmental features, based on L1 norms, largely aligned with what is important for international intelligibility. However, there were divergences in the assessment of suprasegmental features, which did not fully align with the criteria. This is one of the first studies that explored issues of intelligibility directly in the field of language assessment, and provides an alternative way of thinking about the norms and the standard in language assessment from an intelligibility standpoint.	F-Var P-WEng O-Othr TL-Engl
2013	Winters, S., & O'Brien, M. G. (2013). Perceived accentedness and intelligibility: The relative contributions of F0 and duration. *Speech Communication, 55*(3), 486–507.	Similar to some other studies mentioned above (Field, 2005; Hahn, 2004; Zielinski, 2008) examining suprasegmental features affecting L2 intelligibility, **Winters and O'Brien**'s study acoustically investigated the relative contributions of suprasegmental features to perceived accentedness and intelligibility in both L1 and L2 speech. To carry this out, nonnative prosodic features such as syllable duration and F0 contour (intonation) were transferred to native productions, and vice versa. The results of accent rating tasks and sentence transcription tasks indicated that the transfer of prosody across languages had a negative impact on listeners' perceived accentedness and intelligibility. Additionally, nonnative intonation patterns had a more pronounced effect on reducing intelligibility compared to nonnative duration cues. The study offered pedagogical insights into learning prosodic features relevant to intelligibility and accentedness.	F-Var P-Aphon O-Tran TL-Engl TL-Othr

(*Continued*)

(*Continued.*)

Year	References	Annotations	Theme
2014	Derwing, T. M., Munro, M. J., Foote, J. A., Waugh, E., & Fleming, J. (2014). Opening the window on comprehensible pronunciation after 19 years: A workplace training study. *Language Learning, 64*, 526–548. https://doi.org/10.1111/lang.12053	In this study, **Derwing et al.** conducted a pronunciation intervention study for L2 English speakers in the workplace setting with an average time of language immersion for 19 years. Findings suggested that the intervention was effective in improving the comprehensibility and the intelligibility of the speech of the speakers, although no differences were observed in terms of fluency. This study provides results pertaining to an under-researched population, with pedagogical implications targeting this specific L2 speaker group.	F-Var P-Inst P-Pron O-Othr TL-Engl
2014	Jin, S.-H., & Liu, C. (2014). Intelligibility of American English vowels and consonants spoken by international students in the United States. *Journal of Speech, Language, and Hearing Research, 57*(2), 583–596.	Unlike Winters & O'Brien (2013) which mostly focused on suprasegmentals, **Jin and Liu** scrutinized the intelligibility of segmentals (English vowels and consonants) produced by L1 Chinese and Korean speakers. Intelligibility in this study was determined based on native listeners' accuracy in identifying English phonemes pronounced by both native and nonnative speakers. The results indicated that nonnative speakers' vowel intelligibility was significantly lower than their consonant intelligibility to native listeners. Additionally, the study found an association between age of arrival (AOA) and vowel intelligibility, whereas length of residency in the U.S. and frequency of daily English use showed no such relationship. These findings carry pedagogical implications, suggesting training interventions to enhance vowel intelligibility among L2 speakers.	F-Int P-Aphon O-Othr TL-Engl
2014	Kawase, S., Hannah, B., & Wang, Y. (2014). The influence of visual speech information on the intelligibility of English consonants produced by non-native speakers. *The Journal of the Acoustical Society of America, 136*(3), 1352–1362.	**Kawase et al.** investigated how visual (i.e., the speaker's face) and audiovisual (i.e., the speaker's face and voice) speech information affects segmental intelligibility. Using an identification task similar to the method employed by Jin & Liu (2014), they found mixed results: certain features (e.g., /v/, /θ/) were more intelligible with visual information, while others (e.g., /ɹ/) were less intelligible under the same conditions. Compared to previous studies in intelligibility which mainly used audio-only stimuli (e.g., Munro & Derwing, 1995), this study expands the scope of intelligibility research to focus on visual cues which are very much present across various communicative contexts.	F-Var P-Aphon O-Othr TL-Engl
2014	Murphy, J. M. (2014). Intelligible, comprehensible, non-native models in ESL/EFL pronunciation teaching. *System, 42*, 258–269.	In the investigation of intelligible speech, **Murphy** identified various features such as linguistic, paralinguistic, and rhetorical characteristics that contribute to the characterization of intelligible speech. This means that even nonnative speech can be intelligible regardless of accent, as many qualities can influence listeners' intelligibility judgments. By measuring intelligibility through scalar ratings, Murphy's findings extended earlier research (Hayes-Harb & Watzinger-Tharp, 2012; Smith & Rafiqzad, 1979), demonstrating that accent alone does not determine intelligibility. Based on these results, Murphy proposed the possibility of using intelligible nonnative speakers as pronunciation models.	F-Var P-Pron O-Scal TL-Engl
2015	Jułkowska, I. A., & Cebrian, J. (2015). Effects of listener factors and stimulus properties on the intelligibility, comprehensibility and accentedness of L2 speech. *Journal of Second Language Pronunciation, 1*(2), 211–237.	Supplementing previous findings (e.g., Derwing & Munro, 1997; Munro & Derwing, 1995), **Jułkowska and Cebrian** confirmed the partial independence of speech constructs when assessed by three listener groups (native listeners, matched-L1 nonnative listeners, and mismatched-L1 nonnative listeners). However, segmental errors were observed to be associated with accentedness and comprehensibility ratings, but not with those of intelligibility. The findings ultimately supported the intelligibility principle in L2 learning and teaching.	F-SpC F-Var P-Pron O-Tran TL-Engl

(*Continued*)

(*Continued.*)

Year	References	Annotations	Theme
2017	Sheppard, B. E., Elliott, N. C., & Baese-Berk, M. M. (2017). Comprehensibility and intelligibility of international student speech: Comparing perceptions of university EAP instructors and content faculty. *Journal of English for Academic Purposes, 26*, 42–51.	**Sheppard et al.** aimed to investigate potential variances in the intelligibility and comprehensibility of L2 speech to English for Academic Purposes (EAP) instructors versus faculty members teaching content courses. The findings reported no significant difference in transcription accuracy (i.e., intelligibility), but EAP instructors demonstrated greater accuracy in transcribing less intelligible speech compared to content faculty. This finding supports the results of KENNEDY & TROFIMOVICH (2008) and MATSUURA (2007), showing that high intelligibility scores can be affected by listeners' familiarity with target accents.	F-Var P-Pron O-Tran TL-Engl
2018	Kang, O., Thomson, R. I., & Moran, M. (2018). Empirical approaches to measuring the intelligibility of different varieties of English in predicting listener comprehension. *Language Learning, 68*(1), 115–146.	In light of the lack of a standardized intelligibility measure in the field, **Kang et al.** explored various ways of measuring intelligibility by comparing five different intelligibility measures and attempted to identify the most reliable method for assessing L2 speech intelligibility. They examined how segmental and suprasegmental features contribute to intelligibility measures and its relationship with listener comprehension. The results revealed that all phonological variables in the speakers' speech were linked to the intelligibility scores obtained through the scalar rating method. The nonsense sentence measure contributed the most to the listening comprehension test scores. This study not only provided an overview of the operationalization of intelligibility as a speech construct, but also laid the groundwork for future research (e.g., Hirschi & Kang, 2023).	F-Int P-Pron O-Tran O-Othr TL-Engl
2018	Kim, S. (2018). Development of discursive practices for the intelligibility of Thai English in interaction: Sequence and categories as contextual resources. *System, 72*, 164–177.	**Kim**'s (2018) work is particularly interesting because it combines L2 speech intelligibility with innovative analytical methodologies, namely membership categorization analysis (MCA) and conversation analysis (CA). By focusing on the speech of one L2 speaker, the author investigated how the participant utilized contextual resources to produce intelligible L2 pronunciation. Widening the scope of prior research that explored the influence of contextual factors on L2 intelligibility (e.g., GASS & VARONIS, 1984; KENNEDY & TROFIMOVICH, 2008), Kim's study investigated how sequential context and descriptive resources aid in achieving intelligibility during interactions.	F-Var P-Pron O-Tran TL-Engl
2018	Mroz, A. (2018). Seeing how people hear you: French learners experiencing intelligibility through automatic speech recognition. *Foreign Language Annals, 51*(3), 617–637.	This research is influential because **Mroz** explored L2 learners' experiences with intelligibility through ASR technology. The results showed the possibility of ASR being comparable to a native interlocutor's understanding. In addition, L2 learners of French who explicitly learned pronunciation knowledge displayed better attentiveness to the intelligibility of their speech. The study illustrated that the field has begun to consider intelligibility from the perspective of machines, not just from that of human listeners.	F-Int P-Inst P-Tech O-Tran TL-Othr
2020	Kang, O., Moran, M., Ahn, H., & Park, S. (2020a). Proficiency as a mediating variable of intelligibility for different varieties of accents. *Studies in Second Language Acquisition, 42*(2), 471–487.	In this study, **Kang et al.** examined the role of L2 listeners' proficiency in their judgments of comprehension and intelligibility for different types of L2 accented speech. The intelligibility scores were operationalized by a transcription method. The results demonstrated that high-proficiency listeners were not sensitive to accent types when listening to highly intelligible speech, whereas intermediate- to low-level listeners remained sensitive to different accent varieties. This partially supports Kang et al.'s (2019) finding that listeners are not significantly influenced by various accented stimuli if the speech is highly intelligible.	F-Var P-Pron O-Tran TL-Engl

(*Continued*)

(*Continued.*)

Year	References	Annotations	Theme
2020	Kang, O., Thomson, R. I., & Moran, M. (2020b). Which features of accent affect understanding? Exploring the intelligibility threshold of diverse accent varieties. *Applied Linguistics, 41*(4), 453–480.	While many studies have attempted to identify pronunciation features that affect L2 intelligibility (e.g., Field, 2005; Hahn, 2004; Zielinski, 2008), there has not been a clear threshold of intelligibility established (i.e., the minimal number/percentage of L2 features in the speech that would not adversely influence intelligibility). In this regard, **Kang et al.** investigated the phonological features of six English varieties to discover the specific features that determined listeners' understanding of accented speech. Findings reported that vowel/consonant divergence, particularly related to high-functional loads, was the strongest predictor of intelligibility scores, followed by other suprasegmental features such as pace, lexical stress, and tone choices. Moreover, provided that L2 speakers were highly intelligible, their speech had similar intelligibility scores to native speakers, with minimal segmental deviations in content words being a threshold for high intelligibility. This strengthens previous suggestions (Murphy, 2014; Smith & Rafiqzad, 1979) that L2 accented speech can be used in educational or assessment contexts if it is intelligible. Pedagogically, the study informs L2 pronunciation instruction to pinpoint specific speech features that are most likely to impact L2 intelligibility.	F-Var F-Int P-Pron O-Tran TL-Engl
2020	McCrocklin, S., & Edalatishams, I. (2020). Revisiting popular speech recognition software for ESL speech. *TESOL Quarterly, 54*(4), 1086–1097.	In a replication of Derwing et al. (2000), **McCrocklin and Edalatishams** employed a different speech recognition software to assess transcription accuracy and compare it with human listener evaluations. The results mirrored those of the original study, demonstrating higher transcription accuracy for native speech in comparison to nonnative speech. However, unlike Derwing et al. (2000), this study highlighted that the current software's performance closely matched that of human listeners. Their findings also added that ASR recognition was more strongly correlated with human listeners' intelligibility and comprehensibility than with accentedness. These findings imply that L2 teachers can integrate this tool into their classroom practices.	F-Int P-Tech O-Tran TL-Engl
2020	Moussalli, S., & Cardoso, W. (2020). Intelligent personal assistants: Can they understand and be understood by accented L2 learners? *Computer Assisted Language Learning, 33*(8), 865–890.	This study shows the growing focus on AI's intelligibility in L2 research. By employing an intelligent personal assistant (Amazon's Echo), **Moussalli and Cardoso** investigated both the intelligibility of L2-accented speech to AI and the intelligibility of AI-generated speech to L2 learners. Their study operationalized intelligibility (i.e., understanding) through various methods such as ratings, transcriptions, surveys, and interviews, conceptually similar to Kang et al. (2018). The findings unveiled that Amazon's Echo understood L2 accented speech effectively, and L2 learners also exhibited a strong understanding of Echo-generated speech. Pedagogically, this study supports the potential for intelligent personal assistants to serve as valuable teacher assistants in L2 classrooms. Furthermore, this study extends previous research (e.g., Mroz, 2018; 2020), which utilized ASR to investigate intelligibility, by exploring the use of other AI technologies capable of engaging in real-time conversations with L2 speakers.	F-Int P-Tech O-Tran O-Scal O-Othr TL-Engl
2020	Mroz, A. (2020). Aiming for advanced intelligibility and proficiency using mobile ASR. *Journal of Second Language Pronunciation, 6*(1), 12–38.	In this study, **Mroz** aimed to illuminate the effect of ASR technology on L2 learners' proficiency and intelligibility. The study also examined how learner background factors interacted with the use of ASR and how intelligibility related to speaker proficiency, following the intelligibility principle (Levis, 2005). This study was a follow-up to Mroz (2018), extending the investigation to include a comparison between human-based and machine-based ratings. The outcomes revealed a positive improvement in L2 intelligibility and proficiency through the use of ASR, proposing the potential efficacy of integrating ASR into L2 oral and pronunciation teaching and learning methodologies.	F-Int F-Var P-Inst P-Tech O-Tran TL-Othr
2021	Huensch, A., & Nagle, C. (2021). The effect of speaker proficiency on intelligibility, comprehensibility, and accentedness in L2 Spanish: A conceptual replication and extension of Munro and Derwing (1995a). *Language Learning, 71*(3), 626–668.	Building upon the work of Derwing & Munro (1997) and Munro & Derwing (1995), **Huensch and Nagle** not merely replicated, but also extended the interrelations among L2 speech dimensions, with a particular emphasis on intelligibility. The findings indicated that, unlike accentedness, comprehensibility reflected intelligibility more accurately, regardless of speakers' proficiency level. Moreover, since proficiency impacted both linguistic features and listener-based constructs, these linguistic features were found to predict both intelligibility and comprehensibility, but not accentedness. These outcomes have significant pedagogical implications for L2 teaching, emphasizing the importance of comprehensibility as a priority. This approach ultimately leads to improved intelligibility.	F-SpC F-Var P-Pron O-Tran TL-Othr

(*Continued*)

(*Continued.*)

Year	References	Annotations	Theme
2021	Yenkimaleki, M., & Van Heuven, V. J. (2021). Effects of attention to segmental vs. suprasegmental features on the speech intelligibility and comprehensibility of the EFL learners targeting the perception or production-focused practice. *System, 100*(1), 102557.	This study explored a different line of research, examining the benefits of pronunciation training on the intelligibility of EFL learners' speech. It investigated the types of instruction – segmental vs. suprasegmental and production-focused vs. perception-focused – that may improve the intelligibility of L2 speech. **Yenkimaleki and Van Heuven** operationalized intelligibility using a scalar rating system (an 11-point scale from 0). The study recommended that the most effective instruction for enhancing intelligibility for L2 speakers is segmental training, followed by production-focused practice.	F-Var P-Pron O-Scal TL-Engl
2022	Lascotte, D. K., & Tarone, E. (2022). Channeling voices to improve L2 English intelligibility. *Modern Language Journal, 106*(4), 744–763.	Similar to Yenkimaleki and Van Heuven (2021), **Lascotte and Tarone** explored the effect of instruction on the improvement of L2 intelligibility. Specifically, their seven-week instruction involved each L2 learner selecting their own model speaker, either L1 or L2, and imitating their speech as closely as possible, including both verbal and nonverbal patterns. The intelligibility of L2 speakers was rated by experienced ESL instructors. The outcomes of the study indicated that overall, all participants improved their intelligibility, aligning with Murphy's (2014) suggestion of choosing an intelligible nonnative speaker as a model speaker. This finding underscores the need for educators to incorporate more authentic and socially contextualized speaker models in the L2 classroom.	F-Var P-Pron O-Scal TL-Engl
2022	Wheeler, P., & Saito, K. (2022). Second language speech intelligibility revisited: Differential roles of phonological accuracy, visual speech, and iconic gesture. *Modern Language Journal, 106*(2), 429–448.	While most L2 intelligibility research has employed audio-only stimuli, **Wheeler and Saito** explored how various modalities (audio-only, visual speech, and iconic gestures) influenced listeners' intelligibility of L2 speech. This study expanded on the work of Kawase et al. (2014), which examined the role of visual cues in native listeners' understanding of L2-accented speech. However, it differs from previous research by including iconic gestures as a modality and examining the understanding of both L1 and L2 listeners. The study's findings revealed that vowel errors decreased intelligibility for both L1 and L2 listeners, but the presence of iconic gestures significantly enhanced intelligibility. Conversely, visual speech did not have a significant impact on intelligibility. These findings have implications for incorporating nonverbal cues in L2 teaching and learning as learners may benefit from integrating gestures. Furthermore, educators should consider the influence of nonverbal signals on listening and speaking assessments.	F-Var P-Pron O-Tran TL-Engl
2022	Yenkimaleki, M., & Van Heuven, V. J. (2022). Comparing the nativeness vs. intelligibility approach in prosody instruction for developing speaking skills by interpreter trainees: An experimental study. *Speech Communication, 137*, 92–102.	Focusing on instructional approach of intelligibility, **Yenkimaleki and Van Heuven**'s study examined two competing paradigms, i.e., the nativeness and the intelligibility principles in L2 prosody training. Specifically, three groups – the control group, the nativeness-approach group, and the intelligibility-approach group – received their respective training. Then, three experienced raters assessed learners' comprehensibility, accents, and word and sentence stress. While both experimental groups outperformed the control group, the intelligibility-approach group exhibited superior performance compared to the nativeness-approach group across all parameters. Pedagogically, the study emphasizes the positive implications of adopting an intelligibility approach in L2 speaking programs.	F-SpC P-Inst TL-Engl
2023	Ali, M. M. E. (2023). The foreign-accentedness, comprehensibility, and intelligibility of L2 Arabic speech. *Language Teaching Research*, 136,216,882,311,587.	The goal of **Ali**'s study was to gain deeper insights into intelligibility, comprehensibility, and accentedness by utilizing L2 Arabic speech extracted from oral proficiency interviews. Using transcription, these speech samples were found to be highly intelligible to native Arabic listeners. However, they were also rated as highly accented and moderately comprehensible (easy to understand) using the scales. This finding confirms the well-argued claim that accented speech can be highly intelligible (Kang et al., 2020b). In addition, comprehensibility and accentedness were found to be potent predictors of intelligibility, although these three speech constructs exhibited partial independence from each other (Munro & Derwing, 1995). The study offers implications for L2 assessment and instruction.	F-SpC P-Pron O-Tran TL-Othr

(*Continued*)

(*Continued.*)

Year	References	Annotations	Theme
2023	Aoyama, K., Hong, L., Flege, J. E., Akahane-Yamada, R., & Yamada, T. (2023). Relationships between acoustic characteristics and intelligibility scores: A reanalysis of Japanese speakers' productions of American English liquids. *Language and Speech, 66*(4), 1030–1045.	**Aoyama and colleagues** sought to elucidate the relationship between intelligibility and acoustic features of Japanese speakers' productions of American English liquids. Native English-speaking listeners' identification of target stimuli was used to measure intelligibility. This method is similar to the approach used in other studies from the perspective of acoustic phonetics (Jin & Liu, 2014; Kawase et al., 2014). The results reported that not only was the second formant (F2) robustly correlated with intelligibility scores, but also the acoustic parameters of L2 speakers' /l/ and /r/ productions were distinctly different. The study advocates for setting intelligible L2 pronunciation as a realistic goal for L2 learners.	F-Var P-Aphon O-Othr TL-Engl
2023	Hendriks, B., van Meurs, F., & Usmany, N. (2023). The effects of lecturers' non-native accent strength in English on intelligibility and attitudinal evaluations by native and non-native English students. *Language Teaching Research, 27*(6), 1378–1407.	**Hendriks et al.** investigated whether the accent strength (native British English accents, slight or moderate Dutch accents) of L2 speakers could influence intelligibility scores by three listener groups: Dutch listeners, native listeners, and international (non-Dutch/nonnative) listeners. Their measurements included listeners' performance on cloze tasks as well as scalar ratings of comprehensibility and other attitudinal qualities. Consistent with earlier works (e.g., Jułkowska & Cebrian, 2015), all types of speech, regardless of accent strength, were equally intelligible to all listener groups. Additionally, moderately accented speakers received more negative attitudinal evaluations than slightly accented speakers and native speakers. This study underscores the importance of the intelligibility principle and of de-emphasizing native-like proficiency for nonnative speakers of English.	F-Var P-WEng O-Tran TL-Engl
2023	Huensch, A., & Nagle, C. (2023). Revisiting the moderating effect of speaker proficiency on the relationships among intelligibility, comprehensibility, and accentedness in L2 Spanish. *Studies in Second Language Acquisition, 45*(2), 571–585.	Building on their previous research (Huensch & Nagle, 2021), the current study explored the relations among speech constructs and the impact of speaker proficiency on these constructs. Unlike the previous study, this study used a prompted response task. The results remained consistent with earlier findings, as **Huensch and Nagle** observed positive correlations between intelligibility and comprehensibility but no significant correlation between intelligibility and accentedness. This consistency reinforced the partial independence of the three speech constructs (Derwing & Munro, 1997) and suggested that speaking task variations have limited influence on these constructs. Furthermore, the methodological frameworks of Isaacs & Trofimovich (2012) and Kang et al. (2010) have shaped many contemporary studies exploring comprehensibility and intelligibility, including this one.	F-SpC F-Var P-Pron O-Tran TL-Othr
2023	Inceoglu, S., Chen, W.-H., & Lim, H. (2023). Assessment of L2 intelligibility: Comparing L1 listeners and automatic speech recognition. *ReCALL, 35*(1), 89–104.	In line with Mroz (2018, 2020), **Inceoglu et al.** investigated L2 speech intelligibility by operationalizing the construct as the accuracy of transcription for both ASR technology and native listeners. The study targeted vowel minimal pairs and difficult consonants in sentence contexts produced by four Taiwanese EFL learners. The results showed that both rating groups yielded similar word recognition scores. However, for the sentence transcription task, the human listener group outperformed the ASR program. This study underscores the potential of ASR in enhancing L2 pronunciation learning and teaching.	F-Int P-Tech O-Tran TL-Engl
2023	Lee, B. J., & Bailey, J. L. (2023). Assumptions of speaker ethnicity and the effect on ratings of accentedness, comprehensibility, and intelligibility. *Language Awareness, 32*(2), 301–322.	Unlike most studies that assess speech constructs based on accents, **Lee and Bailey**'s research centered on examining the impact of ethnicity on intelligibility, comprehensibility, and accentedness by applying the concept of reverse linguistic stereotyping (RLS) (Kang & Rubin, 2009). In their study, L1 Japanese listeners listened to speech stimuli recorded by native Japanese speakers, paired with photographs depicting individuals of diverse races. Participants perceived nonnative guises as significantly accented compared to native guises, whereas there were no significant differences in their ratings of comprehensibility and intelligibility across different racial guises. This demonstrates that intelligibility is not only partially independent of accent but also partially independent of racial guises.	F-Var P-SocP O-Tran TL-Othr

(*Continued*)

(*Continued.*)

Year	References	Annotations	Theme
2023	Nagle, C. L., Huensch, A., & Zárate, G. (2023). Exploring phonetic predictors of intelligibility, comprehensibility, and foreign accent in L2 Spanish speech. *Modern Language Journal, 107*, 202–221.	In response to the limited empirical evidence regarding the features of intelligible speech, **Nagle et al.** examined the phonetic predictors of three speech constructs in L2 Spanish. In this study, intelligibility was measured by transcription. The findings reported that intelligibility could be predicted by the diphthongization of word-final /o/, velarization of coda /l/, and rising intonation. Interestingly, many other phonetic features did not affect listener understanding significantly, supporting the well-supported notion that accented speech could still be intelligible (Huensch & Nagle, 2023). Furthermore, the results provide useful insights for L2 Spanish pronunciation instructors as they offer a clear understanding of pronunciation features.	F-Var P-Aphon P-Pron O-Tran TL-Othr
2023	Thir, V. (2023). Co-text, context, and listening proficiency as crucial variables in intelligibility among nonnative users of English. *Studies in Second Language Acquisition, 45*(5), 1210–1231.	In this extensive study, **Thir** examined the variables (i.e., co-text, context, familiarity, and listening proficiency) affecting L2 speech intelligibility, operationalized as spoken word recognition (SWR). The study confirmed the significant role of co(n)textual cues in enhancing intelligibility, even more so than speech familiarity. Proficiency strongly influenced how L2 listeners utilized these cues. These findings can be interpreted in light of previous research by Gass & Varonis (1984) and Kennedy & Trofimovich (2008) for co(n)textual cues, and Kang et al. (2020a) for the role of proficiency. Pedagogically, the study advocates for incorporating co(n)textual strategies into L2 pronunciation instruction rather than focusing solely on phonetic accuracy.	F-Var P-WEng O-Tran TL-Engl
2024	Emara, I. F., & Shaker, N. H. (2024). The impact of non-native English speakers' phonological and prosodic features on automatic speech recognition accuracy. *Speech Communication, 157*, 103038.	**Emara and Shaker** aimed to identify the phonological and prosodic features influencing (a) the perception of L2 intelligibility among ESL learners and teachers (see Saito & Van Poeteren, 2012), and (b) the accuracy of ASR. Based on survey responses from teachers and students, mispronunciations of /i/ and /θ/ sounds were identified as the primary phonological features, and speech rate and speech intensity were recognized as prosodic features contributing to intelligible L2 speech. Regarding the ASR results, mispronunciations of /p, θ, ʧ, i, ɔː, ə/ sounds and consonant clusters were prone to intelligibility errors. This study provides useful insights as it demonstrates which pronunciation features affects both human listeners and ASR.	F-Var P-Tech P-Pron O-Tran O-Othr TL-Engl
2024	Hirschi, K., & Kang, O. (2024). Data-driven learning for pronunciation: Perception and production of lexical stress and prominence in academic English. Advance online publication. *TESOL Quarterly*.	Aligning with previous research on the effect of instructions on L2 intelligibility (e.g., Lascotte & Tarone, 2022; Yenkimaleki & Van Heuven, 2021), **Hirschi and Kang** investigated the benefits of Data-Driven Learning for Pronunciation (DDLfP) on improving intelligibility focusing on prominence and lexical stress. The study specifically assessed the impact of the intervention on L2 learners' production using a scale focusing on linguistic features. It was found that DDLfP positively influenced learners' perception and production of prominence and their lexical stress production, although it did not improve their perception of lexical stress. The study suggests that DDLfP could be a potential method to enhance the intelligibility of L2 pronunciation.	F-Var P-Inst P-Pron TL-Engl
2024	Miao, Y., Kang, O., & Meng, X. (2024). Incorporating Global Englishes varieties into EFL classrooms: Development of listening comprehension and pronunciation. Advance online publication. *TESOL Quarterly*.	In this exploratory study, **Miao et al.** implemented an eight-week intervention by introducing and exposing EFL students to diverse English features and explored its effect on students' listening comprehension and pronunciation development across three speech constructs. Intelligibility was determined by accuracy of listener orthographic transcription of short phrases or sentences. Although the intervention improved students' ability to understand different English varieties in a short term, it did not have an effect on students' improvement in any of the speech constructs (accentedness, comprehensibility, and intelligibility). The current findings support the potential usefulness of incorporating diverse English varieties into the English language classroom.	F-Var P-Inst P-WEng O-Tran TL-Engl

[1]Authors' names are shown in small capitals when the study referred to appears elsewhere in this timeline.

Okim Kang is Professor of Applied Linguistics and Director of the Applied Linguistics Speech Lab at Northern Arizona University (NAU), Flagstaff, AZ. Her research interests are speech production and perception, L2 pronunciation and intelligibility, L2 oral assessment and testing, automated scoring and speech recognition, World Englishes, and language attitudes. She has published multiple books, and numerous journal articles and book chapters. She has also delivered over 200 invited talks and conference presentations, and obtained a number of intramural and extramural grants from various federal and testing agencies. She is currently Associate Editor for Applied Linguistics.

Yuna Bae is a Ph.D. student in Applied Linguistics at Northern Arizona University, where she specializes in second language speech research. Her main research interests include speech perception and production, with an emphasis on artificial intelligence (AI) applications, speech technology, language assessment, and prosody in second language acquisition. She is particularly interested in how AI systems perceive and process diverse accented speech and the implications this has for second language learning and teaching.

Yongzhi (Vito) Miao is Assistant Professor of Applied Linguistics at Texas Tech University. His teaching and research interests include second language speech, particularly in relation to: (1) second language acquisition, exploring the inclusion of diverse accent varieties in language classrooms and its impact on students' linguistic development; (2) language assessment, investigating the integration of diverse accents in high-stakes language proficiency tests; (3) research methodology, with a focus on measuring speech constructs as latent variables; and (4) intercultural communication, aiming to facilitate communicative success in English as a lingua franca context by improving listener perception and comprehension of diverse accent varieties.

Cite this article: Kang, O., Bae, Y., & Miao, Y. (2025). L2 speech intelligibility. *Language Teaching*, 1–18. https://doi.org/10.1017/S0261444825101006