15

LISTENING AND COMPREHENSIBILITY

Pavel Trofimovich, Oguzhan Tekin, and Rachael Lindberg

Listeners often find themselves in unexpected, unpredictable, or otherwise nonideal circumstances, such as when they communicate with an interlocutor in a noisy restaurant or experience distorted sound during a video call with a colleague. One experience common to these situations is listeners' effort to understand their interlocutors, and this experience is captured through the construct of comprehensibility, which is the focus of this chapter. Comprehensibility is quintessentially a listener-centered construct, in that it characterizes the subjective experience of ease or difficulty as a listener encounters speech produced by different speakers (e.g., native and nonnative) across a range of listening conditions (e.g., while communicating or only listening) and in various social and professional contexts (e.g., when assessing the speech of a university applicant or engaging in a service encounter at a local business). Comprehensibility is an important construct in language learning and teaching because interaction will succeed as a vehicle for language development only when language learners produce speech that is comprehensible to their interlocutor (Ellis, 2005). Our aim in this chapter is to highlight comprehensibility as a communication-relevant, socially malleable, multidimensional construct that encompasses verbal and nonverbal components, is co-constructed by the listener and the speaker, and has important consequences for interlocutors.

According to Derwing and Munro's influential framework (Derwing & Munro, 2015), which has informed much contemporary work on this construct (for a historical perspective, see Trofimovich et al., 2022), comprehensibility refers to "judgments on a rating scale of how difficult or easy an utterance is to understand" (Derwing & Munro, 1997, p. 2). Comprehensibility is contrasted with intelligibility, which is defined as "the extent to which a speaker's message is actually understood by a listener" (Munro & Derwing, 1995, p. 76). Although related, the two constructs are both conceptually and methodologically distinct. Conceptually speaking, utterances which are unintelligible are generally also low in comprehensibility, but the reverse is not always true as sometimes even perfectly intelligible speech can be quite effortful to listen to and comprehend, for example, because the speaker hesitates a lot or shows little variation in pitch levels, producing monotonous speech (Huensch & Nagle, 2021; Munro & Derwing, 1995;). At the measurement level, comprehensibility is a listener-rated construct, most often captured through Likert-type numerical scales bounded by anchor point descriptors such as "extremely easy to understand" and "impossible to understand" (Munro & Derwing, 1995, p. 79). Conversely, intelligibility is typically operationalized as the accuracy with which listeners orthographically transcribe utterances (Derwing & Munro, 1997) or respond to true–false or comprehension questions after listening to speech content (Hahn, 2004; Kennedy &

DOI: 10.4324/9781003219552-15

Trofimovich, 2008). In essence, comprehensibility is a construct focused on listener understanding of speech, with a strong experiential, subjective processing component.

Given its focus on understanding, comprehensibility is first and foremost useful for researchers and practitioners as a metric of listener understanding. Scalar ratings of comprehensibility are easy to elicit and analyze in various contexts, such as classrooms, workplace settings, assessment situations, and research labs through pre-recorded speech samples or live speaker performances, featuring different length and content. In contrast, measuring intelligibility requires carefully constructed content-specific materials (e.g., audio recordings of target performances, content-relevant comprehension questions). Compared to intelligibility measures, which vary depending on whether they target listeners' understanding of individual words versus their comprehension of discourse-level content (Kang et al., 2018; Kennedy, 2009; 2021), scalar ratings of comprehensibility show high consistency between individual listeners, such that listeners generally agree in their subjective experience of how effortful it is for them to understand a given speech sample (Nagle, 2019; Nagle & Rehman, 2021). Most importantly, comprehensibility ratings provide a usable metric of listeners' actual understanding of speech (Munro & Derwing, 1995; Sheppard et al., 2017).

Through its experiential, subjective component, comprehensibility might also be useful to researchers and practitioners as a measure of processing fluency, which is not the same as the common conceptualization of fluency encompassing temporal aspects of speech production, such as rate and pausing. Processing fluency is instead a metacognitive construct capturing people's perception of the ease or difficulty with which they process information (Reber & Greifeneder, 2017; Schwarz, 2018). Central to research on processing fluency is the idea that a person's subjective experience of ease or difficulty while processing various types of stimuli, such as texts, images, or sounds, might predict their judgment stemming from this experience, including assessments of perceived risk, liking, beauty, ease of learning, and truthfulness (Graf et al., 2018). For example, speakers with lower comprehensibility elicit feelings of annoyance and irritation from listeners, who also judge these speakers as less intelligent and successful compared to more comprehensible speakers (Dragojevic, 2020). Comprehensibility is thus an appealing measure of processing fluency, with potential implications for listeners' reactions, and ultimately their behaviors, toward speakers.

Critical Issues and Topics

The main strands of second language (L2) comprehensibility research can be summarized under several perspectives, collectively describing this construct as having strong speaker- and listener-based contributions, as a variable with visual and behavioral components, as a dynamic, multidimensional, and interaction-relevant phenomenon, and as a socially flexible judgment with important attitudinal and behavioral consequences.

Speaker Perspective

One key issue underlying the construct of comprehensibility is the extent to which comprehensibility assessments reflect the linguistic content of speech being evaluated by listeners. If language teachers and learners consider comprehensible speech, rather than nativelike, non-accented L2 performance, to be a worthwhile goal of L2 pronunciation learning and teaching (Levis, 2020), then it is important to know which linguistic dimensions make L2 speech more or less comprehensible to listeners. According to a recent meta-analysis of 37 studies focusing on L2 English (Saito, 2021), listener-rated comprehensibility is broadly associated with several linguistic features, including those that capture a speaker's segmental production (e.g., accuracy of individual vowels and consonants), prosody (e.g., word stress placement, intonation accuracy), and temporal fluency (e.g., speech rate, pausing), in addition to multiple measures of lexis, grammar, and discourse, such as the variety and richness of vocabulary, accuracy and complexity of morphosyntax, and discourse organization

(Appel et al., 2019; Isaacs & Trofimovich, 2012; Saito et al., 2016). This finding is robust, attested across several languages, including French (Bergeron & Trofimovich, 2017), German (O'Brien, 2014), and Japanese (Saito & Akiyama, 2017). Although the importance of specific linguistic dimensions relevant to comprehensibility may differ depending on the type of speaking performance being evaluated, for example, a picture description versus a source-based academic presentation (Crowther, 2020; Crowther et al., 2018), language teachers and learners might welcome the news that comprehensibility is not only about pronunciation but can be achieved by targeting other aspects of language, including grammar and vocabulary. This creates interesting challenges for language teaching and assessment, where grammar and vocabulary are typically taught and evaluated separately from pronunciation. What emerges here, broadly speaking, is a need to reconceptualize and integrate various skills, including grammar, vocabulary, and various aspects of pronunciation, in teaching and assessment (Isaacs et al., 2018).

Listener Perspective

Because comprehensibility is a listener-centered construct, a critical question is whether comprehensibility judgments depend on who is judging comprehensibility. There is now a substantial knowledge base suggesting that comprehensibility judgments vary as a function of various listener characteristics, including linguistic training (Isaacs & Thomson, 2013), familiarity with the target language (Munro et al., 2006), language teaching and learning experience (Saito et al., 2017; 2019), knowledge of and proficiency in multiple languages (Saito & Shintani, 2016), and awareness of the importance of comprehensible speech for communication (Saito et al., 2019). Researchers have also expanded their listener sample, recruiting not only native speakers of the target language but also L2 speakers (Crowther et al., 2016; O'Brien, 2014), bilinguals and multilinguals (Saito & Shintani, 2016), and members of specific professional communities (Derwing & Munro, 2009; Kennedy & Trofimovich, 2014; Sheppard et al., 2017). Despite various differences across listeners in their linguistic, social, and professional backgrounds, according to a recent meta-analysis (Saito, 2021), listeners demonstrate high consistency in their comprehensibility ratings, regardless of whether they are trained or novice raters, or L2 speakers themselves (Crowther et al., 2016; Derwing & Munro, 2013; Saito et al., 2017).

Even though different listeners generally agree in their ratings of comprehensibility, the stated reasons for their assessments often diverge (Foote & Trofimovich, 2018; Isaacs & Thomson, 2020). For example, when language teachers were asked to comment on their ratings of L2 French speakers, the teachers commented on various linguistic dimensions in the speakers' speech, but each highlighted a different issue (e.g., quality of pronunciation, accuracy of individual vowels and consonants) that was particularly relevant to their ratings (Kennedy et al., 2017). In another study, even though French- and Mandarin-speaking listener groups did not differ in their assessments, the French listeners brought up a more diverse set of linguistic dimensions that impacted their judgments, citing issues of pronunciation, fluency, lexis, and grammar, compared to the Mandarin listeners who generally reported on fluency (Crowther et al., 2016). Thus, even though different listeners, including native- and L2-speaking raters, generally achieve high rating consistency, providing similar evaluations, they may arrive at comparable judgments by relying on different criteria.

While listener-specific differences might not matter for comprehensibility assessment in many informal or low-stakes contexts (e.g., when speaking with a barista in a café or communicating with a peer student in a group activity), these differences could take on particular significance in higher-stakes situations where a speaker's comprehensibility might contribute to decision-making, for example, regarding a student's course grade, an applicant's university admission, or a candidate's job suitability. There are several ways in which individuals who manage assessment can minimize inter-individual differences across listeners. For example, Staples et al. (2014) engaged listeners in

informal contact activities during which they collaborated with L2 speakers, and those who participated in these activities showed reduced severity in their evaluations of comprehensibility. In another study, Kermad (~~2021~~) engaged inexperienced, novice listeners in a two-hour instructional intervention, introducing them to the construct of comprehensibility and providing them with varied samples of L2 speaking performances, with further opportunities for the listeners to practice and compare their ratings. The intervention resulted in reduced rater severity, increased rater precision, and improved rater consistency, all enhancing the validity of comprehensibility assessments. These research insights are complemented by work in language assessment, for instance, when speaking examiners receive training to develop "an international ear," so that they can minimize the impact of various differences across listeners on their assessments (Taylor & Galaczi, 2011).

Visual and Behavioral Perspective

In most experimental settings, comprehensibility has been evaluated through listener ratings of audio recordings. However, in many operational contexts, L2 speakers are judged when a rater (listener or interlocutor) also has access to visual information. This raises an important question of whether a listener's perception of comprehensibility is influenced by speakers' visual cues (e.g., eyebrow raises, blinks, hand gestures) and their behaviors such as providing backchannels (e.g., *uh huh*, *yeah*, *right*), displaying visual signals of understanding (e.g., nodding), or demonstrating reciprocity and mutuality in conversation (e.g., by encouraging the partner to speak through nonverbal cues or elaborating the partner's idea). It might well be that the ease or difficulty with which listeners understand an L2 speaker is shaped not just by what they hear but also by what they observe.

With respect to the role of visual cues in listener assessments of comprehensibility, Tsunemoto, Lindberg, et al. (2022) showed that, with more visual information, listeners tend to be less severe in their ratings, judging L2 speakers easier to understand. Speakers elicited the lowest ratings when they were assessed through audio recordings only. When listeners had access to video clips showing the speakers' facial expressions but not gestures, the ratings improved. The same speakers received the highest ratings when their facial expressions and gestures were fully visible in unedited video clips. These initial findings, which are consistent with work targeting visual cues in listening comprehension (Batty, 2015; Wagner, 2008), imply that comprehensibility may be shaped by visual information, all of which may help listeners anticipate upcoming information, thereby decreasing their processing effort.

In terms of speakers' behaviors and their potential impact on comprehensibility, Trofimovich et al. (2021) analyzed 36 paired oral interactions involving university-level L2 speakers in an academic discussion task, coding speaker behaviors for eight measures of task engagement, including cognitive engagement (idea units, language-related episodes), social engagement (encouragement, responsiveness, task and time management, backchanneling, nodding), and emotional engagement (positive affect through laughing and smiling). Four measures showed associations with comprehensibility, as rated by each speaker's partner at the end of the task. The speakers who encouraged their partners to claim a turn or sought a response from them and those who provided signs of acknowledgment through nodding were perceived as easier to understand. However, the speakers who produced more frequent language-related episodes (discussing language forms or engaging in self- or peer-correction) and demonstrated more responsiveness (essentially repeating, completing, or commenting on previously expressed ideas) were rated as harder to understand. The current work on visual, nonverbal, and behavioral contributions to comprehensibility, which is compatible with research on interactional competence (Galaczi & Taylor, 2018; see also Chapter 24 "Listening and Interactional Competence," in this volume), is clearly in its infancy, and further focused research is needed to inform researchers and practitioners about the various verbal and nonverbal features making speakers more or less comprehensible to listeners.

Dynamic and Multidimensional Perspective

One key question is whether and how listeners' decision-making about a speaker's comprehensibility evolves as they gain more experience with speech. A related question is whether comprehensibility is co-constructed by listeners, especially when it is assessed in conversation. Put differently, comprehensibility might be the outcome of a dynamic process whereby listeners process varying levels of accuracy, complexity, and fluency in a speaker's speech over time, and listeners (as interlocutors) might draw on various verbal and nonverbal means to continuously shape what they consider to be comprehensible. In an initial project targeting these questions, Nagle et al. (2019) investigated listeners' judgments of comprehensibility in real time, as they evaluated 3-minute audio-recorded speaking performances by L2 Spanish speakers through a computer interface, where the listeners could increase or decrease their rating at particular points in the narrative. The listeners displayed several distinct rating patterns, ranging from highly dynamic (demonstrating large fluctuations) to non-dynamic (showing only a handful, smaller-scale increases or decreases), highlighting comprehensibility as a person-specific and time-varying construct. Several linguistic dimensions in the speakers' speech (discourse organization, fluency) emerged as positive attractors, leading to upgrades in comprehensibility, while other dimensions (lexical and grammatical errors) seemed to decrease comprehensibility.

In a follow-up study, Trofimovich et al. (2020) reasoned that interactive speech, where interlocutors co-construct discourse in real time, might be even more revealing of the time-sensitive, evolving nature of comprehensibility, compared to a one-way listening task. They analyzed 17-minute collaborative conversations between 40 L2 English university students from different language backgrounds, where the two speakers rated each other's comprehensibility seven times at 2–3-minute intervals. The speakers' comprehensibility ratings demonstrated a U-shaped function, with comprehensibility rated initially high, then downgraded in the middle of the interaction, and finally reaching high levels again at the end of the conversation. Moreover, after the first few minutes, the speakers became aligned in how they perceived each other's comprehensibility, revealing a rapid adjustment in how interlocutors co-construct each other's comprehensibility and implying that (with sufficient time or increased interpersonal comfort) comprehensibility issues might become less severe or noticeable for both interactants. While the speakers were nearly always aligned in their perception of each other's comprehensibility, when external listeners evaluated the same speaking performances through audio recordings, their ratings were significantly lower than the speakers' assessments in six of the seven rating episodes, meaning that the external listeners found it more effortful to understand the speakers than the interactants themselves (Nagle et al., 2022b). Speaker comprehensibility might therefore depend on whether those who assess it are active participants in the conversation. Whereas interaction partners can deploy various tools at their disposal to ensure mutual understanding (e.g., giving and receiving verbal and nonverbal cues, adjusting speech rate and the use of pauses, predicting forthcoming utterances), external listeners do not have access to many of these resources as they assess the amount of effort required to understand the speech.

Nagle et al. (2022a) subsequently revisited the same dataset, integrating the speakers' comprehensibility ratings with their self- and peer-evaluations of anxiety and collaborativeness on the assumption that comprehensibility might reflect interpersonal dimensions of interaction such as affect and behavior. The speakers' comprehensibility was predicted by how anxious the speakers were judged to be by their interlocutors, where lower perceived anxiety was associated with greater comprehensibility, and also by how collaborative they seemed to their interlocutors, such that greater perceived collaborativeness was linked to more comprehensible speech. Moreover, the speakers' self-rated collaborativeness also positively predicted how comprehensible they sounded to their interlocutor, suggesting that comprehensibility, at least to some degree, also reflected the speakers' own perception of how much or how little they contributed to the interaction. Thus, it appears that comprehensibility is dynamic, co-constructed, and multidimensional, stemming from all interlocutors in a dialogue.

Social Perspective

Speech assessment, including listener-based evaluations of comprehensibility, takes place in a social context, where listeners could not only develop impressions about speakers, but also overtly or subtly express attitudinal judgments about those speakers arising from listeners' specific prior experiences or internalized beliefs (Dragojevic et al., 2021). Considering that comprehensibility assessments are common in many low- and high-stakes contexts (Jones & Isaacs, 2021), it is therefore important to understand how listeners' attitudes, among other influences, such as biases, stereotypes, and priming effects, influence their ratings and whether these influences can be mitigated.

There is a growing body of work suggesting that listener-based comprehensibility assessments are subject to social influences (Taylor Reid et al., 2022). For example, when university faculty were asked to rate speech samples recorded by international students, the faculty who expressed negative judgments about the language skills of international students perceived the students as harder to understand than the faculty whose attitudes were more positive, even though both groups did not differ in their actual understanding of the speech content (Sheppard et al., 2017). In another study, before asking listeners to evaluate comprehensibility, Taylor Reid et al. (2019) presented listeners with either a positive or a negative story about the language skills of L2 speakers. Relative to the assessments of baseline listeners (who heard no biasing story), the same L2 speakers were evaluated more favorably by positively biased listeners, whereas they were rated less favorably by negatively biased listeners, particularly older persons who likely had stronger, well-established social beliefs. Taylor Reid, O'Brien, Trofimovich, and Bajt (2020) replicated these findings with teachers, where native-speaking teachers of L2 German were affected by negative bias, assigning lower comprehensibility ratings to L2 speakers, compared to the teachers who heard no biasing story.

However, social bias in comprehensibility ratings can be mitigated through various interventions (Kang et al., 2015). For instance, Taylor Reid et al. (2022) engaged listeners in perspective taking, which they defined as asking listeners to perform the same task as the speakers to be assessed, on the assumption that task practice might minimize or eliminate social influences on speech ratings (see also Taylor Reid, O'Brien, Trofimovich, & Tsunemoto, 2020). Those who engaged in task practice indeed demonstrated reduced negative rating bias, likely because of their enhanced familiarity with the task and finer-grained understanding of performance expectations. These findings highlight comprehensibility as a socially relevant construct driven by social forces, and call for future work focused on identifying and minimizing social bias in listener evaluations of L2 speakers.

Recommendations for Practice

The good news is that comprehensibility can be improved through instruction. Based on recent meta-analytic evidence, targeted instruction results in sizeable gains in L2 comprehensibility, as indicated through comparisons of pre- and post-instruction ratings assessed by various listeners, including expert judges and untrained raters (Saito, 2021; Saito & Plonsky, 2019). More importantly, consistent with the idea that listeners draw on multiple linguistic dimensions to evaluate a speaker's comprehensibility, various types of instructional approaches appear effective (Saito, 2021), including instruction focusing on speech prosody, pronunciation of individual vowels and consonants, and various fluency phenomena. Comprehensibility might also develop through extensive interactive experiences, as learners engage in face-to-face communication with a partner or via online video tools (Saito & Akiyama, 2017), or through sustained communicative teaching over multiple months, often without a dedicated pronunciation focus (French et al., 2020; Nagle, 2018). To improve their students' comprehensibility, instructors might additionally target specific techniques, such as shadowing activities, where students mimic what they hear in an audio or video recording soon after the original speaker (Foote & McDonough, 2017), imitation practice (Ding et al., 2019), and drama-based tasks (Galante & Thomson, 2017). Thus, language teachers and learners should approach comprehensibility as a

construct that encompasses multiple elements, not just pronunciation, targeting comprehensibility through different types of experiences which can involve work on vocabulary, grammar, fluency, or discourse organization. Teachers might also highlight links between comprehensible speech and speakers' anxiety and their interactive, collaborative behaviors, and point to visual cues enhancing a speaker's comprehensibility for listeners. Above all, learners should be encouraged to engage in guided analysis of their own and other speakers' comprehensibility (Tsunemoto, Trofimovich, et al., 2022), alternating between the speaker and the listener perspectives, so they can develop an understanding of how comprehensibility is co-constructed by speakers and listeners.

Given that comprehensibility assessment is subject to various social influences, researchers and practitioners might wish to consider ways in which attitudes or biases might impact listener evaluations of L2 speech in various educational and professional settings. Awareness-raising, perspective-taking, diversity training, or informal-contact activities might be useful, depending on the context and situation-specific constraints (Kang & Moran, 2019; Kim et al., 2019). As an example of perspective-taking, before they begin their assessments, language teachers and test examiners could be asked to complete tasks similar to those they are assessing and to establish or refresh their understanding of performance benchmarks. Similarly, in research or professional settings, individual listeners or evaluation committees might read other people's narratives or share their own stories about experiencing prejudice or empathy from their interlocutors based on their linguistic performance. In terms of contact activities, for instance, in workplaces with a significant multilingual, multicultural workforce, formal or informal activities, or language classes can also be implemented. In sum, various forms of awareness-raising and extended contact opportunities might have the double benefit of enhancing comprehensibility for speakers and listeners and mitigating potential social biases in listener-based assessments.

Future Directions

Comprehensibility-focused work is a vibrant area of research activity as it deals with a construct that is not only complex and multidimensional but one with ample practical implications. While researchers might be interested in uncovering linguistic, social, experiential, affective, metacognitive, and behavioral dimensions of comprehensibility, language teachers and learners might wish to work on comprehensible oral production or develop better listening skills. For example, researchers and practitioners would benefit from a clearer understanding of the links between comprehensibility and various measures of understanding, including intelligibility and listening comprehension (e.g., listener-based transcription, responses to comprehension questions, story retell). While comprehensibility ratings frequently overlap with more objectively measured understanding (Munro & Derwing, 1995), the magnitude of this link might vary for different speakers and listeners (Matsuura et al., 1999). In fact, the extent of overlap between measures of listening comprehension and comprehensibility for the same group of listeners evaluating a given audiovisual material, such as an academic lecture or an unscripted conversation, is largely unknown, because few studies include both sets of measures in the same dataset and those that do rarely compare them (Kang et al., 2018; Sheppard et al., 2017).

Given the lack of longitudinal research, researchers could also engage in multistage projects exploring the interplay between L2 speakers' cognitive, motivational, experiential, and affective profiles in the development of comprehensible L2 speech in different instructed and uninstructed settings. Future work could also center on the listener perspective by comparing listeners who participate in versus those who observe speaker performances or by targeting listeners with distinct individual profiles, such as those with musical training or with high tolerance for ambiguity or noise. Researchers could also elaborate on various facets of listener experience, including how long it takes for listeners to process speech (Ludwig & Mora, 2017) or how irritable they feel (Ludwig, 1982), and

could identify multiple behavioral correlates of comprehensibility such as listeners' ability to shadow L2 speech (Inoue et al., 2018).

With the view of refining and extending a focus on interactive, nonverbal, and behavioral components of comprehensible L2 speech, researchers might also wish to intensify work on interaction-driven comprehensibility in various specific contexts (Pavlenko et al., 2019; Tsunemoto, McAndrews, et al., ~~2022~~). Given that interlocutor comprehensibility in interaction might be co-constructed and tightly coordinated, it might be interesting to explore the validity of a joint (rather than speaker-specific) measure of comprehensibility for both interlocutors in a conversation. Last but not least, researchers could continue examining the role of speakers' sociopolitical views, stereotypical judgments, and attitudes toward the speaker or the topic of conversation in their comprehensibility judgments, with the view of developing and testing the effectiveness of different bias mitigation strategies.

Finally, researchers might wish to embrace the issue raised by Varonis and Gass (1982): Why do listeners react to L2 speakers in particular ways? In their classic study, Varonis and Gass observed listeners respond to L2 speakers asking for directions. The listeners tended to repeat the request (often with a rising intonation), showing some reluctance to get involved in a conversation, often accompanied by a sigh or a filler like *oh boy*, which was interpreted as a direct reaction to the speaker's comprehensibility. Even though the listeners fully understood the speaker's message, their experience was effortful and the likelihood of a future nonunderstanding was real, so the listeners' behaviors likely arose as a consequence of speaker comprehensibility.

As should become obvious from this chapter, comprehensibility has generally been investigated as the target and the end goal of research, with researchers examining various speaker and listener influences on comprehensibility. In fact, apart from research on processing fluency, whose goal is to explain listeners' affective and attitudinal reactions to speakers as a function of listening effort (Dragojevic, 2020; Jensen & Thøgersen, 2020), comprehensibility has rarely been framed in the sense intended by Varonis and Gass, namely, as a *predictor* of people's reactions and behaviors.

Therefore, considering the point raised by Varonis and Gass (1982), and consistent with the processing fluency perspective, it might be useful for researchers to adopt comprehensibility as an explanatory variable, extending its use as a predictor of a greater range of listener behaviors, beyond affective and attitudinal judgments, so that researchers and practitioners could gain an understanding of the processing costs and rewards of listener experience with low- versus high-comprehensibility speech. Our preliminary work has begun to explore these issues with respect to L2 speakers' perception of their interactive experience (Nagle et al., 2023). This work draws on a large dataset involving 90 pairs of L2 speakers engaged in paired interaction in three collaborative tasks over 30 minutes, where all interacting partners evaluate their own and each other's comprehensibility three times during conversation and then provide overall assessments of their experience at the end of the session. The speakers' self-rating of comprehensibility emerged as a significant predictor of their overall experience and of their communication success (see also Tekin et al., 2022). These preliminary findings suggest that comprehensibility explains, at least to some degree, how L2 speakers evaluate the success of their interactive experience. These results also imply a speaker-centric role of comprehensibility in interaction, in the sense that what ultimately matters for communication success, as perceived by its participants, might stem from the perceptions of the speakers themselves, rather than determined by the judgments of their interlocutors.

Conclusion

Taken broadly, the research findings reviewed here highlight comprehensibility as an exciting and worthwhile target of study, with implications for language speakers in various contexts. Comprehensibility is relevant to speakers who would like to develop better communication skills and to listeners who, as professional raters or casual observers, evaluate speakers' linguistic or professional competences. The current research findings can already speak to some aspects of these

real-world concerns. With a view to the future, however, comprehensibility might also help explain various human judgments and behaviors, for instance, whether interlocutors' willingness to communicate with a speaker fluctuates as a function of the speaker's comprehensibility or whether the perceived effectiveness of paired-interactive activities for L2 learners depends on the comprehensibility of the speaker they are paired with. However, these issues (and undoubtedly many others) await further exploration by new generations of researchers, who will refine and certainly transform our understanding of what it means for language speakers to produce and perceive comprehensible L2 speech.

Further Reading

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

This book features a state-of-the-art review of literature on various constructs relevant to L2 pronunciation, including comprehensibility. This research-informed volume, written in an engaging, clear language, also provides an insightful discussion of how various errors contribute to comprehension difficulty and outlines various approaches to the teaching and assessment of pronunciation, including comprehensibility. The volume can be used as both a textbook in a university-level course for students interested in second language teaching and learning and as engaging standalone reading for layperson readers.

Isaacs, T., & Trofimovich, P. (Eds.). (2016). *Second language pronunciation assessment: Interdisciplinary perspectives*. Multilingual Matters.

Written in an accessible language and addressed to a wide range of readers, including researchers, language teachers, and teacher educators, this open-access edited volume consists of five main parts and features multiple research contributions providing a holistic look at pronunciation assessment, including the assessment of comprehensibility. In line with the complex and interdisciplinary nature of pronunciation assessment, this volume not only draws on the important work in psycholinguistics, sociolinguistics, and second language acquisition among others, but also brings together research on various skills, such as listening and writing, to promote future research on pronunciation assessment.

Isaacs, T., Trofimovich, P., & Foote, J. A. (2018). Developing a user-oriented second language comprehensibility scale for English-medium universities. *Language Testing*, *35*, 193–216.

Comprehensibility plays an important role in pronunciation assessment. However, few practical tools are available to assess this construct outside the research context, contributing to a gap between theory and practice. With this in mind, the researchers in this study conducted nine focus group interviews with English for Academic Purposes teachers in the UK and Canada to develop and validate an easy-to-use comprehensibility-focused scale. The developed scale, which is freely accessible through online supporting documentation, can be used for formative assessments of university-level students (pre- and post-admission) from a variety of language backgrounds.

References

Appel, R., Trofimovich, P., Saito, K., Isaacs, T., & Webb, S. (2019). Lexical aspects of comprehensibility and nativeness from the perspective of native-speaking English raters. ITL–International *Journal of Applied Linguistics*, *170*(1), 24–52. https://doi.org/10.1075/itl.17026.app

Batty, A. O. (2015). A comparison of video- and audio-mediated listening tests with many-facet Rasch modeling and differential distractor functioning. *Language Testing*, *32*(1), 3–20. https://doi.org/10.1177/0265532214531254

Bergeron, A., & Trofimovich, P. (2017). Linguistic dimensions of accentenedness and comprehensibility: Exploring task and listener effects in second language French. *Foreign Language Annals*, *50*(3), 547–566. https://doi.org/10.1111/flan.12285

Crowther, D. (2020). Rating L2 speaker comprehensibility on monologic versus interactive tasks: What is the effect of speaking task type? *Journal of Second Language Pronunciation*, *6*(1), 96–121. https://doi.org/10.1075/jslp.19019.cro

Crowther, D., Trofimovich, P., & Isaacs, T. (2016). Linguistic dimensions of second language accent and comprehensibility: Nonnative listeners' perspectives. *Journal of Second Language Pronunciation*, *2*(2), 160–182. https://doi.org/10.1075/jslp.2.2.02cro

Crowther, D., Trofimovich, P., Saito, K., & Isaacs, T. (2018). Linguistic dimensions of L2 accentedness and comprehensibility vary across speaking tasks. *Studies in Second Language Acquisition*, *40*(2), 443–457. https://doi.org/10.1017/S027226311700016X

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition*, *19*(1), 1–16. https://doi.org/10.1017/S0272263197001010

Derwing, T. M., & Munro, M. J. (2009). Comprehensibility as a factor in listener interaction preferences: Implications for the workplace. *Canadian Modern Language Review*, *66*(2), 181–202. https://doi.org/10.3138/cmlr.66.2.181

Derwing, T. M., & Munro, M. J. (2013). The development of L2 oral language skills in two L1 groups: A 7-year study. *Language Learning*, *63*(2), 163–185. https://doi.org/10.1111/lang.12000

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

Ding, S., Liberatore, C., Sonsaat, S., Lučić, I., Silpachai, A., Zhao, G, … Gutierrez-Osuna, R. (2019). Golden speaker builder: An interactive tool for pronunciation training. *Speech Communication*, *115*, 51–66. https://doi.org/10.1016/j.specom.2019.10.005

Dragojevic, M. (2020). Extending the fluency principle: Factors that increase listeners' processing fluency positively bias their language attitudes. *Communication Monographs*, *87*(2), 158–178. https://doi.org/10.1080/03637751.2019.1663543

Dragojevic, M., Fasoli, F., Cramer, J., & Rakić, T. (2021). Toward a century of language attitudes research: Looking back and moving forward. *Journal of Language and Social Psychology*, *40*(1), 60–79. https://doi.org/10.117/0261927X20966714

Ellis, R. (2005). Principles of instructed language learning. *System*, *33*(2), 209–224. https://doi.org/10.1016/j.system.2004.12.006

Foote, J. A., & McDonough, K. (2017). Using shadowing with mobile technology to improve L2 pronunciation. *Journal of Second Language Pronunciation*, *3*(1), 34–56. https://doi.org/10.1075/jslp.3.1.02foo

Foote, J., & Trofimovich, P. (2018). Is it because of my language background? A study of language background influence on comprehensibility judgments. *Canadian Modern Language Review*, *74*(2), 253–278. https://doi.org/10.3138/cmlr.2017-0011

French, L. M., Gagné, N., & Collins, L. (2020). Long-term effects of intensive instruction on fluency, comprehensibility and accentedness. *Journal of Second Language Pronunciation*, *6*(3), 380–401. https://doi.org/10.1075/jslp.20026.fre

Galaczi, E. D., & Taylor, L. (2018). Interactional competence: Conceptualizations, operationalizations, and outstanding questions. *Language Assessment Quarterly*, *15*(3), 219–236. https://doi.org/10.1080/15434303.2018.1453816

Galante, A., & Thomson, R. I. (2017). The effectiveness of drama as an instructional approach for the development of second language oral fluency, comprehensibility, and accentedness. *TESOL Quarterly*, *51*(1), 115–142. https://doi.org/10.1002/tesq.290

Graf, L. K. M., Mayer, S., & Landwehr, J. R. (2018). Measuring processing fluency: One versus five items. *Journal of Consumer Psychology*, *28*(3), 393–411. https://doi.org/10.1002/jcpy.1021

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, *38*(2), 201–223. https://doi.org/10.2307/3588378

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*(2), 626–668. https://doi.org/10.1111/lang.12451

Inoue, Y., Kabashima, S., Saito, D., Minematsu, N., Kanamura, K., Yamauchi, Y. (2018). A study of objective measurement of comprehensibility through native speakers' shadowing of learners' utterances. *Proceedings of Interspeech*, *2018*(2), 1651–1655. https://doi.org/10.21437/Interspeech.2018-1860

Isaacs, T., & Thomson, R. I. (2013). Rater experience, rating scale length, and judgments of L2 pronunciation: Revisiting research conventions. *Language Assessment Quarterly*, *10*(2), 135–159. https://doi.org/10.1080/15434303.2013.769545

Isaacs, T., & Thomson, R. I. (2020). Reactions to second language speech: Influences of discrete speech characteristics, rater experience, and speaker first language background. *Journal of Second Language Pronunciation*, *6*(2), 402–429. https://doi.org/10.1075/jslp.20018.isa

Isaacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility: Identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition*, *34*(2), 475–505. https://doi.org/10.1017/S0272263112000150

Isaacs, T., Trofimovich, P., & Foote, J. A. (2018). Developing a user-oriented second language comprehensibility scale for English-medium universities. *Language Testing*, *35*, 193–216.

Jensen, C., & Thøgersen, J. (2020). Comprehensibility, lecture recall and attitudes in EMI. *Journal of English for Academic Purposes*, *48*, Article 100912. https://doi.org/10.1016/j.jeap.2020.100912

Jones, J., & Isaacs, T. (2021). Assessing second language pronunciation. In H. Moehebbi & C. Coombe (Eds.), *Research questions in language education and applied linguistics* (pp. 305–310). Springer.

Kang, O., & Moran, M. (2019). Enhancing communication between ITAs and US undergraduate students. In S. D. Looney & S. Bhalla (Eds.), *A transdisciplinary approach to international teaching assistants: Perspectives from applied linguistics* (pp. 82–99). Multilingual Matters.

Kang, O., Rubin, D., & Lindemann, S. (2015). Mitigating US undergraduates' attitudes toward international teaching assistants. *TESOL Quarterly*, *49*(2), 681–706. https://doi.org/10.1002/tesq.192

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*(2), 115–146. https://doi.org/10.1111/lang.12270

Kennedy, S. (2009). L2 proficiency: Measuring the intelligibility of words and extended speech. In A. Benati (Ed.), *Issues in second language proficiency* (pp. 132–144). Continuum.

Kennedy, S. (2021). Difficulties understanding L2 speech due to discourse- versus word-level elements. *Journal of Second Language Pronunciation*, *7*(3), 315–342. https://doi.org/10.1075/jslp.18015.ken

Kennedy, S., Blanchet, J., & Guénette, D. (2017). Teacher-raters' assessment of French lingua franca pronunciation In T. Isaacs & P. Trofimovich (Eds.), *Second language pronunciation assessment: Interdisciplinary perspectives* (pp. 210–236). Multilingual Matters.

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), 459–490. https://doi.org/10.3138/cmlr.64.3.459

Kennedy, S., & Trofimovich, P. (2014). First-and final-semester non-native students in an English-medium university: Judgments of their speech by university peers. *Language Learning in Higher Education*, *3*(2), 283–303. https://doi.org/10.1515/cercles-2013-0015

Kermad, A. (2021). Training the "everyday" listener how to rate accented speech. *International Journal of Listening*. ~~Published online November 12, 2021.~~ https://doi.org/10.1080/10904018.2021.1987910

Kim, R., Roberson, L., Russo, M., & Briganti, P. (2019). Language diversity, nonnative accents, and their consequences at the workplace: Recommendations for individuals, teams, and organizations. *The Journal of Applied Behavioral Science*, *55*(1), 73–95. https://doi.org/10.1177/0021886318800997

Levis, J. (2020). Revisiting the intelligibility and nativeness principles. *Journal of Second Language Pronunciation*, *6*(3), 310–328. https://doi.org/10.1075/jslp.20050.lev

Ludwig, A., & Mora, J. C. (2017). Processing time and comprehensibility judgments in non-native listeners' perception of L2 speech. *Journal of Second Language Pronunciation*, *3*(2), 167–198. https://doi.org/10.1075/jslp.3.2.01lud

Ludwig, J. (1982). Native-speaker judgments of second-language learners' efforts at communication: A review. *The Modern Language Journal*, *66*(3), 274–283. https://doi.org/10.1111/j.1540-4781.1982.tb06989.x

Matsuura, H., Chiba, R., & Fujieda, M. (1999). Intelligibility and comprehensibility of American and Irish Englishes in Japan. *World Englishes*, *18*(2), 49–62. https://doi.org/10.1111/1467-971X.00121

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. https://doi.org/10.1111/j.1467-1770.1995.tb00963.x

Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition*, *28*(1), 113–131. https://doi.org/10.1017/S0272263106060049

Nagle, C. (2018). Motivation, comprehensibility, and accentedness in L2 Spanish: Investigating motivation as a time-varying predictor of pronunciation development. *The Modern Language Journal*, *102*(1), 199–217. https://doi.org/10.1111/modl.12461

Nagle, C. (2019). Developing and validating a methodology for crowdsourcing L2 speech ratings in Amazon Mechanical Turk. *Journal of Second Language Pronunciation*, *5*(2), 294–323. https://doi.org/10.1075/jslp.18016.nag

Nagle, C., & Rehman, I. (2021). Doing L2 speech research online: Why and how to collect online ratings data. *Studies in Second Language Acquisition*, *43*(4), 916–939. https://doi.org/10.1017/S0272263121000292

Nagle, C., Trofimovich, P., & Bergeron, A. (2019). Toward a dynamic view of second language comprehensibility. *Studies in Second Language Acquisition*, *41*(4), 647–672. https://doi.org/10.1017/S0272263119000044

Nagle, C., Trofimovich, P., O'Brien, M. G., & Kennedy, S. (2022a). Beyond linguistic features: Exploring the behavioral and affective correlates of comprehensible second language speech. *Studies in Second Language Acquisition*, *44*(1), 255–270. https://doi.org/10.1017/S0272263121000073

Nagle, C., Trofimovich, P., O'Brien, M. G., & Kennedy, S. (2022b). Comprehensible to whom? Examining rater, speaker, and interlocutor perspectives on comprehensibility in an interactive context. *The Modern Language Journal*, ~~Advance online publication.~~ https://doi.org/10.1111/modl.12809

Nagle, C., Trofimovich, P., Tekin, O., & McDonough, K. (2023). Framing second language comprehensibility: Do interlocutors' ratings predict their perceived communicative experience? *Applied Psycholinguistics*, *44*(1), 131-156. https://doi.org/10.1017/S0142716423000073

O'Brien, M. G. (2014). L2 learners' assessments of accentedness, fluency, and comprehensibility of native and nonnative German speech. *Language Learning*, *64*(4), 715–748. https://doi.org/10.1111/lang.12082

Pavlenko, A., Hepford, E., & Jarvis, S. (2019). An illusion of understanding: How native and non-native speakers of English understand (and misunderstand) their Miranda rights. *International Journal of Speech, Language, and the Law*, *26*(2), 181–207. https://doi.org/10.1558/ijsll.39163

Reber, R., & Greifeneder, R. (2017). Processing fluency in education: How metacognitive feelings shape learning, belief formation, and affect. *Educational Psychologist*, *52*(2), 84–103. https://doi.org/10.1080/00461520.2016.1258173

Saito, K. (2021). What characterizes comprehensible and native-like pronunciation among English-as-a-second-language speakers? Meta-analyses of phonological, rater, and instructional factors. *TESOL Quarterly*, *55*(3), 866–900. https://doi.org/10.1002/tesq.3027

Saito, K., & Akiyama, Y. (2017). Linguistic correlates of comprehensibility in second language Japanese speech. *Journal of Second Language Pronunciation*, *3*(2), 199–217. https://doi.org/10.1075/jslp.3.2.02sai

Saito, K., & Plonsky, L. (2019). Effects of second language pronunciation teaching revisited: A proposed measurement framework and meta-analysis. *Language Learning*, *69*(3), 652–708. https://doi.org/10.1111/lang.12345

Saito, K., & Shintani, N. (2016). Do native speakers of North American and Singapore English differentially perceive comprehensibility in second language speech? *TESOL Quarterly*, *50*(2), 421–446. https://doi.org/10.1002/tesq.234

Saito, K., Tran, M., Suzukida, Y., Sun, H., Magne, V., & Ilkan, M. (2019). How do L2 listeners perceive the comprehensibility of foreign-accented speech? Roles of L1 profiles, L2 proficiency, age, experience, familiarity and metacognition. *Studies in Second Language Acquisition*, *41*(5), 1133–1149. https://doi.org/10.1017/S0272263119000226

Saito, K., Trofimovich, P., & Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics*, *37*(2), 217–240. https://doi.org/10.1017/S0142716414000502

Saito, K., Trofimovich, P., Isaacs, T., & Webb, S. (2017). Re-examining phonological and lexical correlates of second language comprehensibility: The role of rater experience. In T. Isaacs & P. Trofimovich (Eds.), *Second language pronunciation assessment: Interdisciplinary perspectives* (pp. 141–156). Multilingual Matters.

Schwarz, N. (2018). Of fluency, beauty, and truth: Inferences from metacognitive experiences. In J. Proust & M. Fortier (Eds.), *Metacognitive diversity: An interdisciplinary approach* (pp. 25–46). Oxford University Press.

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*(1), 42–51. https://doi.org/10.1016/j.jeap.2017.01.006

Staples, S., Kang, O., & Wittner, E. (2014). Considering interlocutors in university discourse communities: Impacting US undergraduates' perceptions of ITAs through a structured contact program. *English for Specific Purposes*, *35*(1), 54–65. https://doi.org/10.1016/j.esp.2014.02.002

Taylor, L., & Galaczi, E. (2011). Scoring validity. In L. Taylor (Ed.), *Examining speaking: Research and practice in assessing second language speaking* (pp. 171–233). Cambridge University Press.

Taylor Reid, K., O'Brien, M. G., Trofimovich, P., & Bajt, A. (2020). Testing the malleability of teachers' judgments of second language speech. *Journal of Second Language Pronunciation*, *6*(2), 236–264. https://doi.org/10.1075/jslp.19015.tay

Taylor Reid, K., O'Brien, M. G., Trofimovich, P., & Tsunemoto, A. (2020). Exploring the stability of second language speech ratings through task practice in bilinguals' two languages. *Journal of Monolingual and Bilingual Speech*, *2*(2), 315–329. https://doi.org/10.1558/jmbs.17714

Taylor Reid, K., Trofimovich, P., & O'Brien, M. G. (2019). Social attitudes and speech ratings: Effects of positive and negative bias on multiage listeners' judgments of second language speech. *Studies in Second Language Acquisition*, *41*(2), 419–442. https://doi.org/10.1017/S0272263118000244

Taylor Reid, K., Trofimovich, P., & O'Brien, M. G. (2022). The malleability of listener judgments of second language speech. In V. Sardegna & A. Jarosz (Eds.), *Theoretical and practical developments in English speech assessment, research, and training* (pp. 11–26). Springer. https://doi.org/10.1007/978-3-030-98218-8_2

Taylor Reid, K., Trofimovich, P., O'Brien, M. G., & Tsunemoto, A. (2022). Using task practice to reduce social influences on listener evaluations of second language accent and comprehensibility. *International Journal of Listening*, *22*(3), 283–298. https://doi.org/10.1080/10904018.2021.1904933

Tekin, O., Trofimovich, P., Chen, T.-H., & McDonough, K. (2022). Alignment in second language speakers' perceptions of interaction and its relationship to perceived communicative success. *System*, *108*, Article 102848. https://doi.org/10.1016/j.system.2022.102848

Trofimovich, P., Isaacs, T., Kennedy, S., & Tsunemoto, A. (2022). Speech comprehensibility. In T. M. Derwing, M. J. Munro, & R. Thomson (Eds.), *The Routledge handbook of second language acquisition and speaking* (pp. 174–187). Routledge.

Trofimovich, P., Nagle, C. L., O'Brien, M. G., Kennedy, S., Taylor Reid, K., & Strachan, L. (2020). Second language comprehensibility as a dynamic construct. *Journal of Second Language Pronunciation*, *6*(3), 430–457. https://doi.org/10.1075/jslp.20003.tro

Trofimovich, P., Tekin, O., & McDonough, K. (2021). Task engagement and comprehensibility in interaction: Moving from what second language speakers say to what they do. *Journal of Second Language Pronunciation*, *7*(3), 435–461. https://doi.org/10.1075/jslp.21006.tro

Tsunemoto, A., Lindberg, R., Trofimovich, P., & McDonough, K. (2022). Visual cues and rater perceptions of second language comprehensibility, accentedness, and fluency. *Studies in Second Language Acquisition*, *44*(3), 659–684. https://doi.org/10.1017/S0272263121000425

Tsunemoto, A., McAndrews, M., Trofimovich, P., & Friginal, E. (~~2022~~). Listener perceptions of customer service agents' performance: Do comprehensibility, accentedness, and fluency matter? *Journal of Second Language Pronunciation*. ~~Advanced online publication.~~ https://doi.org/10.1075/jslp.21027.tsu

Tsunemoto, A., Trofimovich, P., Blanchet, J., Bertrand, J., & Kennedy, S. (2022). Effects of benchmarking and peer-assessment on French learners' self-assessments of accentedness, comprehensibility, and fluency. *Foreign Language Annals*, *55*(1), 135–154. https://doi.org/10.1111/flan.12571

Varonis, E. M., & Gass, S. (1982). The comprehensibility of non-native speech. *Studies in Second Language Acquisition*, *4*(2), 114–136. https://doi.org/10.1017/S027226310000437X

Wagner, E. (2008). Video listening tests: What are they measuring? *Language Assessment Quarterly*, *5*(3), 218–243. https://doi.org/10.1080/15434300802213015