OPEN ACCESS

Australian Journal of Applied Linguistics

ISSN 2209-0959

https://www.castledown.com/journals/ajal/

Australian Journal of Applied Linguistics, 6(1), 40–54 (2023)
https://doi.org/10.29140/ajal.v6n1.1026

Fluency, comprehensibility, and accentedness in L2 speech: Examining the role of visual and acoustic information in listener judgments

BURCU GOKGOZ-KURT [a]

[a] *Kutahya Dumlupinar University, TÜRKIYE*
burcugokgozkurt@fulbrightmail.org

Abstract

The purpose of the present study was to examine (a) the effects of audio-visual information and (b) the role of temporal measurements associated with the perceptual ratings of fluency, accentedness, comprehensibility, pronunciation, and oral proficiency in second language (L2) speech samples of International Teaching Assistants (ITAs). American undergraduate students randomly assigned to audiovisual and audio-only (without seeing the speaker) conditions rated the two-minute speech samples of five ITAs on a seven-point scale. For acoustic analysis, several temporal measurements were used that were hypothesized to have an association with the ratings. Findings revealed that listeners without visual information judged L2 speech as less fluent and more accented although no significant difference in comprehensibility was found. L2 speakers were also judged to have better pronunciation and higher oral proficiency when the recordings were presented in video mode. The findings further indicated that among the acoustic measures examined, articulation rate was found to be associated with perceptual ratings. Implications for ITA training and language testing and assessment are discussed.

Keywords: accentedness, acoustic analysis, comprehensibility, visual cues, fluency

Introduction

The use of visual/nonverbal cues by second language (L2) speakers, such as the movements related to face, hands, head, or posture, has been shown to enhance L2 listeners' speech perception (e.g., Buck, 2001; Wheeler & Saito, 2022) and elevate their ability to comprehend the conveyed message (e.g., Dahl & Ludvigsen, 2014; Sueyoshi & Hardison, 2005). These findings indicate that when L2 speech is accompanied by visual information, it is likely to create a more positive impression on L2 listeners (Tsunemoto et al., 2022). This, in turn, influences listeners' judgments of L2 speech in various global

Data Availability Statement: *All relevant data are in this paper.*

constructs, including spoken fluency (henceforth fluency), comprehensibility (ease of understanding L2 speech), and accentedness (an approximation to nativelike speech) as indicators of L2 speakers' oral proficiency (Derwing & Munro, 2015). However, previous studies have more often investigated these constructs by having listeners rate the audio-only recordings (Trofimovich et al., 2020; for an exception see Tsunemoto et al., 2022), which does not reflect the authentic mode of communication. Thus, the primary goal of the present study is to address this gap by investigating the impact of presentation mode (audio vs. audiovisual) on the ratings of fluency, comprehensibility, and accentedness by listeners who speak English as their first language (L1).

All three constructs have been extensively studied by employing listeners' scalar ratings as well as several linguistic dimensions, including acoustic measures. Research has indicated that examining the acoustic measures of L2 speech will contribute to a more comprehensive understanding of L2 speech features (De Jong et al., 2012; Tavakoli et al., 2020). Thus, utilizing a combination of listener judgments (*perceived fluency*) and objective speech measures (*utterance fluency*) is still desirable (Segalowitz, 2010, pp. 29-31, 2016). This holds for a better understanding of comprehensibility and accentedness, as well, given that various fluency measures (e.g., speech rate) were found to be associated with accentedness (e.g., Kang, 2012; Kang et al., 2010; Saito et al., 2017b) and comprehensibility (e.g., Crowther et al., 2018; Saito et al., 2017b). However, studies are yet to examine similar relationships among these constructs of L2 speech under varying rating conditions, including those with and without the visual input of the speaker.

This is especially critical for understanding whether and how the presence of visual information affects the inexperienced L1 listeners' perceptual judgments of L2 speakers, who are expected to make themselves understood in most educational and professional contexts. University college campuses in countries where English is spoken as an L1 constitute an example of these since they host international students from around the world. International teaching assistants (ITAs) are defined as international graduate students responsible for teaching various undergraduate-level classes at colleges mostly in countries such as the U.S.A. or Canada (Gorsuch, 2016). Since ITAs are typically L2 speakers, universities ensure that, before hiring them for an undergraduate teaching position, their proficiency in written and spoken English is sufficient. This is mainly because ITAs are expected to communicate effectively with students in the classrooms or labs. However, this does not prevent them from having to face listener bias (Kang & Rubin, 2009; Lindemann & Subtirelu, 2013) such as incomprehensibility (Munro & Derwing, 1995) or stereotyping (Lindemann, 2003). Due to such bias, ITAs may be perceived as less efficient as instructors compared to their native-English-speaking peers, as evidenced in college students' relatively low teacher evaluations[1] (Kang, et al., 2015; Meyer & Mao, 2014; see also Sanchez & Khan, 2016). Thus, it is essential to identify the factors that are linked to the perceptual assessments of ITA speech by inexperienced listeners, specifically college students in this study, to enhance communication in college classrooms. In this regard, it would be useful to obtain a better understanding of the effects of visual information on listeners' ratings of ITAs' L2 speech and their acoustic correlates. As such, the present study sets out to examine (a) whether and how judgments of fluency, accentedness, comprehensibility, pronunciation, and oral proficiency change when raters see the speakers in the recordings (audiovisual condition) as opposed to when they do not (audio condition), and (b) how objective fluency measures and listener ratings are related.

Review of Literature

Fluency, comprehensibility, and accentedness in L2 speech

The following section provides an examination of L2 fluency, comprehensibility, and accentedness, along with relevant research related to these constructs. It is crucial to present and explore each of these constructs, and subsequently, discuss pertinent research findings to support each definition. To begin with, fluency, the primary construct being investigated, can be defined as a "smoothly and effortlessly functioning psycholinguistic process" (Lennon, 1990, p. 391) although there is a lack of agreement on how to use the term in the literature (Chambers, 1997; Derwing et al., 2004; Wood, 2001; for a review, see Kormos, 2006; Tavakoli & Wright, 2020). Schmidt (1992), for example, differentiates between fluent and nonfluent speakers based on the differences in false starts, pauses,

and other signs of hesitation (p. 377), indicating that temporal aspects are important in predicting perceived fluency. In fact, scholars already established a connection between various temporal measures of L2 speech such as speech rate (e.g., Derwing et al., 2004; Rossiter, 2009), mean length of runs, phonation-time ratio, pace (Kormos & Dénes, 2004) and listener judgments, whereas the variance was not consistent across studies as reported by the meta-analysis conducted by Suzuki et al. (2021). More data coming from different listener and speaker groups will shed more light on the type and structure of this link.

The critical position of fluency in L2 oral proficiency assessments is apparent in the leading British- and American-based high-stakes English proficiency tests such as the Test of English as a Foreign Language (TOEFL) of Educational Testing Service (ETS) and the International English Language Testing System (IELTS). For example, one of the four categories in the rubric used for the independent speaking task in the TOEFL (ETS, 2004) is called "delivery" which requires speakers to have "[a] generally well-paced flow (fluid expression)" during speech performance. Similarly, in the IELTS, a speaker in the highest band is described as "[someone who] speaks fluently with only rare repetition or self-correction; [and] any hesitation is content-related rather than to find words or grammar."[2] Despite the subjectivity and variation in the interpretation of the term fluency across tests (for a discussion, see De Jong, 2018), this emphasis on fluency indicates a better understanding of fluency is also crucial for achieving academic and professional success by ITAs (Gorsuch, 2011).

The next construct investigated in this study is comprehensibility, which is widely associated with understanding the meaning of an utterance. By definition, comprehensibility is the listener's perception and assessment of the amount of effort needed to understand an utterance (Munro & Derwing, 1999, 2001; for a discussion, see Levis, 2006). However, previously, it was not considered to be very different from the concept of intelligibility (Gass & Varonis, 1984), which usually refers to what is understood in an utterance, whereas comprehensibility is regarded as "the perception of intelligibility" (Derwing & Munro, 1997, p. 2). According to Isaacs and Trofimovich (2011), having a certain degree of fluency is a must for holding comprehensibility scores of "mid- to upper-range." High-stakes proficiency tests also rely on comprehensibility judgments of speakers as their assessments are based on listeners' perceptions of how difficult or easy a speech performance is to understand (Saito et al., 2017b). For these reasons, comprehensibility rather than intelligibility has been the subject of inquiry in this study.

Finally, the construct of accentedness refers to a listener's assessment of a speaker's foreign accent strength or "linguistic nativelikeness" (Saito et al., 2016). Given that the primary purpose of L2 communication is being understood, accentedness has not been prioritized in L2 classrooms. However, it has been widely studied in relation to comprehensibility (Derwing & Munro, 1997, 2005), though less so to fluency. Studies have demonstrated that accentedness and comprehensibility are considered to be "partially overlapping but essentially independent" constructs (Saito & Akiyama, 2017, p. 2; see also Derwing & Munro, 2005; Kang et al., 2010; but see Huensch & Nagle, 2021), but there seems to be more "overlap" in other domains such as pronunciation and fluency (Saito et al., 2017b).

Given the definitions of each construct, there is a need for an examination of research establishing the links among these concepts. Since college students constitute the target audience for ITAs and are "relatively reliable" in their ratings (Derwing et al., 2004; Rossiter, 2009), further studies investigating their perceptual judgments could very much contribute to the existing literature. Research indicated that L1 listeners were more severe in their judgments of accentedness than in their ratings of comprehensibility (Derwing & Munro, 1997; Saito et al., 2016). Similarly, Derwing et al. (2004) found that comprehensibility was more correlated with fluency ratings than accentedness. Their results suggested a weak relationship between accentedness and fluency, especially in low-proficiency speakers, which is in line with what Burgess (2001) observed for advanced-level L2 speakers. Another study by Pinget et al. (2014) looked at L2 Dutch learners' accentedness and fluency as rated by L1 Dutch speakers and found that objective measures of both constructs were related to ratings. Finally, a series of related studies (e.g., Kang, 2012; Kang et al., 2010) investigated how measures of speech rate, pausing, stress, and intonation affect judgments of comprehensibility and accentedness by American undergraduate students. The findings reveal that measures such as mean length of runs, articulation rate, and phonation-time ratio help predict about one-third of the variance in comprehen-

sibility judgments, while lower fluency ratings meant lower accentedness judgments.

Therefore, as previous research has shown, comprehensibility and accentedness were often associated with certain fluency measures (e.g., Crowther et al., 2018; Kang, 2012; Kang et al., 2010; Saito et al., 2017b), which may be primarily because fluency is a linguistic dimension based on which raters perform (Isaacs & Trofimovich, 2011). On the other hand, accentedness appears to have a weaker association with fluency and comprehensibility ratings, and it is generally evaluated more strictly. Nevertheless, it remains to be explored how presentation mode, i.e., with and without visual input and acoustic measurements of L2 speech are related to the ratings of fluency, comprehensibility, and accentedness.

The influence of visual information on the ratings of L2 speech

The present study also explores the effects of visual input on the listeners' ratings of L2 speech. A speech presented in audio mode allows listeners to concentrate on the content without being distracted by visuals, especially in situations where listeners are only required to comprehend the speech rather than interact with an interlocutor (Buck, 2001). In line with this view, it has been suggested that the role of nonverbal information, specifically gestures, is "redundant" in communication (Krauss et al., 1991, p. 752). On the other hand, visual information has also been found to be especially helpful for listeners because it carries cues expressed in various ways such as body language, facial expressions, or lip movements (Buck, 2001; Rubin, 1995). Evidence from L2 learners also suggests that processing a message is difficult when a speaker is not present as a part of the spoken text (Hasan, 2000; Thompson & Rubin, 1996). These visual cues have been referred to as "co-text" which is essential for making sense of the message (Rost, 2011; see also Fukumura et al., 2010). Supporting the "bimodal and bisensory" model of speech perception and its essential role in communication, Harris (2003) maintains that "the presentation of dialogue in an audiovisual mode, as opposed to an audio-only mode, preserves important con-textual features of conversation, like gestures, that may facilitate L2 listening comprehension" (p. 184). It should, however, be noted that seeing the speakers rather than just listening to them may also cause stereotyping of speakers and lead to less favorable judgments of accentedness as shown by studies using guise procedures (see Rubin, 1992).

Using nonverbal cues in speech was also shown to be crucial especially for "borderline [ITA] test takers" in their oral interviews as assessed by trained raters, although no such advantage was found for "linguistically proficient" ITAs (Jenkins & Parra, 2003). In their recent study, Nakatsuhara and her colleagues (2021) examined how live, audio, and video modes affected raters while they scored IELTS speaking test performances. The findings indicated that raters noted more negative aspects in non-live rating sessions and that the audio ratings were lower than the ones in the video mode, which was in support of previous research (Conlan et al., 1994). This provides evidence for some raters' considering various non-linguistic cues while rating. On the other hand, the rater reliability was claimed to be higher in ratings using audio recordings because, in the video mode, raters were shown to be distracted by visual cues. Thus, there was more variability in their judgments as they were less likely to focus on the linguistic elements in the performances. In terms of comprehension, studies have already shown that presenting materials using video as opposed to audio mode promotes better understanding (Burgoon et al., 2016) although the video was shown to cause increased attention demand for the raters (Bejar et al., 2000; for a discussion, see Nakatsuhara et al., 2021). Finally, a recent study by Tsunemoto et al. (2022) also looked at the role of visual cues, specifically facial expressions and hand gestures, in L2 speech assessment. Sixty university students with different L1 backgrounds rated 20 speech samples of Chinese and Spanish speakers for fluency, comprehensibility, and accentedness on a 1000-point scale, and the findings indicated that raters with more visual cues were likely to perceive the speaker as more comprehensible and less accented than those with fewer visual cues. However, in contrast to this study, the raters in their study were undergraduate and graduate students with diverse L1 backgrounds.

The Current Study

Despite the growing interest in the literature discussed so far, which examines the constructs of fluency, comprehensibility, and accentedness, whether and how the audio-visual, as opposed to audio

presentation mode, affects listeners' ratings of these constructs (e.g., Tsunemoto et al., 2022) and how temporal variables are associated with these ratings were seldom explored. Thus, the present study seeks to investigate the impact of visual information (seeing the speaker) on listeners' ratings of fluency, accentedness, and comprehensibility as well as pronunciation and (global) oral proficiency ratings. It also looks at whether and to what extent the temporal variables explain the ratings of these constructs. The present study thus grounds for an exploration of these issues to address the following research questions:

1. Whether and how does the presentation mode (audiovisual vs. audio-only) affect the ratings of fluency, comprehensibility, accentedness, as well as pronunciation, and oral proficiency in ITA speech?
2. What relation exists between various temporal measures of ITA speech samples (specifically, articulation rate, phonation-time ratio, mean length of runs, and silent pausing) and listeners' perceptions of fluency, comprehension, accent, pronunciation, and oral proficiency?

Methodology

Speech samples

The ITA speech samples selected for the study were drawn from a database of ITA candidates' micro -lesson presentations recorded for educational purposes. Specifically, the speakers were video-recorded in a classroom setting during the ITA oral proficiency exam required of all ITA candidates wishing to teach classes in their academic departments (a valid speaking subscore of 26+ in the TOEFL IBT or 7+ in the IELTS was required for exemption). After ethical clearance, speech performances by five speakers were selected from the database based on certain criteria. First, speech performances were matched for topic complexity and similarity. Some variety in L1 background and L2 proficiency was also considered since the purpose of the study was not to measure listener bias from a social or psychological perspective. For the audiovisual presentation mode, to control for the visual information variable, dynamic waist shots (from the waist up) were used (see Drijvers & Özyürek, 2017, 2019), and the frequency and duration of gesture use were maximally matched across five speakers. The resolution quality and distance of the speakers from the camera made it difficult to discern their facial expressions.

Raters

The study participants included 31 freshmen enrolled in their second term at a large research university in the southeastern United States. They were recruited through convenience sampling on a voluntary basis. The participants performed the ratings and filled out a background questionnaire in a classroom setting in two intact groups: audio-visual group (AVG; M_{age}= 20.2, SD = 1.08) and audio group (AG; M_{age}= 19.8 years, SD = .97). In the AVG, 6 were female and 8 were male, and in the AG, 8 were female and 9 were male. Nine students (n_{AVG} = 4; n_{VG} = 5) had lived abroad for more than 15 days while none of the students, except one in the AVG group, had any extensive contact with L2 English speakers. Of 20 students (n_{AVG} = 8; n_{VG} = 12), who previously had an instructor speaking L2

Table 1 *ITAs' L1 background and oral proficiency in L2 English*

Speakers	Native Language	Oral Proficiency Score* (Max.100)
Speaker 1	Mandarin	85
Speaker 2	Hindi	83
Speaker 3	Hindi	76
Speaker 4	Mandarin	71
Speaker 5	Mandarin	68

Note *The assessment rubric is based on pronunciation (with the highest proportion), grammar, vocabulary, fluency, and linguistic and extralinguistic awareness.

English in their content classes, 10 students (n_{AVG} = 4; n_{VG} = 6) described their experience as being "OK", 7 (n_{AVG} = 3; n_{VG} = 4) being "Negative", and 3 (n_{AVG} = 1; n_{VG} = 2) being "Positive." On average, participants had somewhat infrequent communication with L2 English speakers when asked to rate their typical weekly contact with them on a scale out of 5. (M_{AVG} = 2.26; M_{AG} = 2.29). No students reported hearing or sight problems. Besides, since the present did not aim at understanding how listener variables help explain the ratings, for controlling listener variability, the responses of one student who was a bilingual speaker of Spanish and English were not included in the analysis (N = 31).

Procedures

Rating procedure: Fluency, comprehensibility, and accentedness

There are two groups of freshmen involved in the rating procedure. Once the informed consent was obtained, the participants completed a language background questionnaire. While the first group of students (*n* = 14) watched the ITAs' video recordings, the second group (*n* = 17) only listened to their audio recordings. There were five recordings per group, each of which lasted for two minutes. Raters in two intact groups were randomly assigned to audio-visual or audio groups through cluster random sampling with a quasi-experimental design. The rating scale was distributed to raters, and no additional explanations other than those written on the scale were provided. The researcher informed that they would listen to or watch a speaker at a time, depending on the group they were assigned to, and rate each using the rating scale. Speech samples were presented over loudspeakers connected to a laptop computer in a carpeted classroom setting. The participants judged each sample for comprehensibility, accentedness, and fluency, as well as pronunciation and oral proficiency using a 7-point scale operationalized following previous research (Kang, 2010; Kennedy & Trofimovich, 2008; Munro & Derwing, 1999). The rating scale included two items per construct (fluency, comprehensibility, and accentedness), the scores of which were then averaged. There was only one item to measure pronunciation and oral proficiency. Also, a score of "1" represented the lowest score on the scale (e.g., comprehensibility, 1 = *hard to understand*, 7 = *easy to understand*). Once the rating procedure was completed, the questionnaire with sociodemographic and language background questions was distributed to the raters.

Acoustic analysis: Temporal measures

Besides the perceptual judgments, an acoustic analysis of five speech samples was performed in *Praat 5.3.4.47* (Boersma & Weenink, 2013) using a script by De Jong and Wempe (2009). It measures speech and articulation rate by detecting syllable nuclei with intensity (dB) and voicedness. Potential syllable nuclei are determined by "peaks in intensity (dB) that are preceded and followed by dips in intensity" (De Jong & Wempe, 2009). Based on this measure, the number of syllables, silent pauses, and speech time were calculated by the author and were inspected manually by one trained speech analyst for the calculations of articulation rate, phonation-time ratio, mean length of runs, and the number and duration of silent pauses. Previous research cautioned against the use of "internally related" or overlapping measures of fluency in making conclusions about utterance fluency (Bosker et al., 2013; Kormos, 2006; Suzuki et al., 2021; Tavakoli, 2016; but see De Jong, 2018). Therefore, a combination of speed and breakdown fluency measures operationalized following the established indices were used in the present study (De Jong & Wempe, 2009; Tavakoli & Wright, 2020, pp. 53-56).

Data Analysis

First, because the data was continuous, the interrater reliability index for each measure was computed using Cronbach's alpha, initially across all 31 raters, and then across raters in each group, as summarized in Table 2.

Raters showed strong reliability across all rated variables (a = .94-.98). Therefore, the perceptual ratings provided for each construct were averaged to calculate the mean scores. Table 3 summarizes the descriptive statistics for mean scores. Assumptions associated with the normality of the distribu-

Table 2 *Interrater reliability for ratings of speech samples*

Rated Variable	All Raters (N = 31)	Audio-visual Group (AVG) (n = 14)	Audio Group (AG) (n = 17)
Fluency	.97	.96	.95
Comprehensibility	.98	.97	.97
Accentedness	.98	.97	.97
Pronunciation	.98	.96	.97
Oral Proficiency	.97	.94	.96

tions were examined, and no significant violations were noted except for one value belonging to "Accentedness" in the AG. Furthermore, homogeneity of variance also appeared plausible looking at similar SDs. The data distribution for each group was sufficiently normal for conducting a t-test as shown by the values in Table 3, as well as nonsignificant Shapiro-Wilk results for each construct (p =.2 – p =.8). The assumption of homogeneity of variances was also tested and met via Levene's F test, with values ranging from $F(29) = 2.19$ $p = .15$ to $F(29) = .01$ $p = .9$.

Regarding acoustic analyses, due to the small sample size, Kendall's tau-b, a non-parametric statistical test was selected to correlate the perceptual ratings with acoustic measures to see how each temporal measure is related to the ratings. For manual calculations of utterance fluency, intercoder reliability was calculated using Cohen's kappa and the result was high at .87.

Findings

Comparison of perceptual ratings by AVG and AG listeners

The first objective of this study was to investigate the perceptions of fluency, comprehensibility, accentedness, pronunciation, and oral proficiency under two different presentation modes (AVG and AG) by two groups of native English-speaking listeners. The analysis aimed at examining the relationship among five rated variables across groups. The descriptive statistics regarding each rated measure are presented in Table 3.

Descriptive statistics show that the mean ratings by the listeners in AVG were higher in all rated variables, which indicates that visual information affects listeners positively in their ratings of L2 fluency. Figure 1 visualizes the effect of audiovisual presentation mode in ratings displaying the spread of the effect in each group.

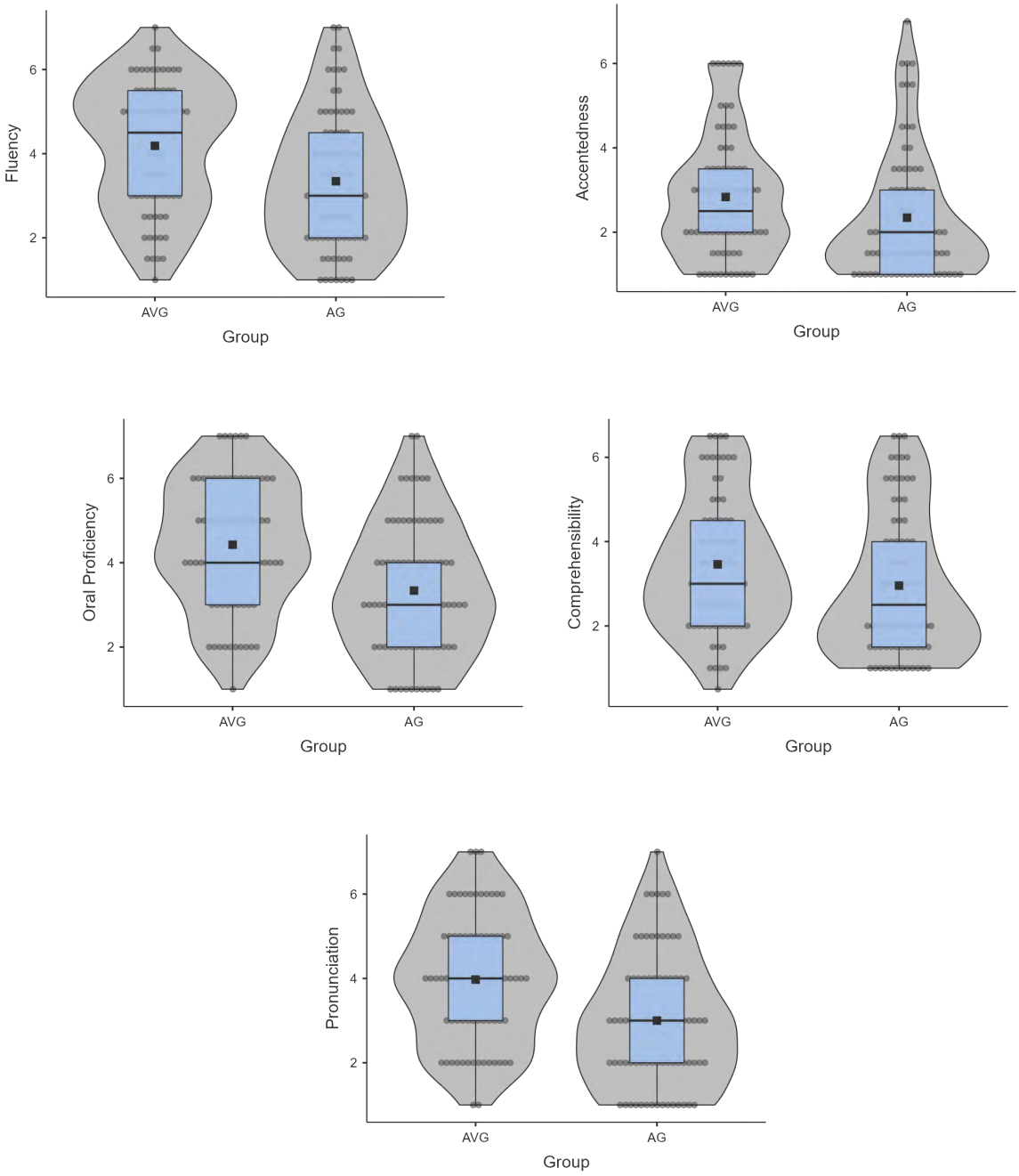
To see whether these differences in mean scores are meaningful, an independent samples t-test was conducted, and it was found to be associated with a statistically significant effect for all five constructs. The results of the analysis are presented in Table 4 below.

Table 3 *Descriptive statistics of ratings in two rating modes*

Groups	Audiovisual Group (n = 14)				Audio Group (n = 17)			
Variables	M*	SD	Ske	Kurt	M*	SD	Ske	Kurt
Comprehensibility	3.46	1.58	0.38	-0.75	2.96	1.68	0.67	-0.80
Accentedness	2.83	1.44	0.72	-0.17	2.34	1.48	1.32	1.06
Fluency	4.19	1.50	-0.29	-1.01	3.34	1.59	0.44	-0.60
Pronunciation	3.98	1.52	0.07	-0.82	3.00	1.54	0.51	-0.52
Oral Proficiency	4.43	1.54	-0.16	-0.84	3.33	1.56	0.34	-0.58

Note. *The maximum score was 7.

Figure 1 *Violin plots displaying the ratings by AVG and AG in five measurement products*



Note. The lines surrounding the box plots are called density curves or kernel density plots. At the center of each density curve, there is a small box plot that displays the first and third quartiles' boundaries with a rectangle and the median with a central dot.

By plotting these cell means, the results revealed that the difference between two groups in ratings of five different measurement products depended on the raters' being in AG or AVG with the raters in the latter having significantly higher ratings in four of the five constructs, which is also illustrated in Figure 1. Except for comprehensibility ratings (p = .047), all other rated variables were found to be significant (p = .047). However, significant group effects were found for fluency [$t(29)$ = 3.03, p = .005, d = 1.09], pronunciation [$t(29)$ = 3.96, p = .001, d = 1.422], and oral proficiency [$t(29)$ = 3.49, p = .002, d = 1.256], all with large effect size, and accentedness with a medium to large effect size [$t(29)$ = 2.31, p = .028, d = .830] (Plonsky & Oswald, 2014).

Table 4 *Independent samples t-test summary table*

	T	*p*	Mean difference	SE difference	Effect Size**	95% CI*** Lower	Upper
Comprehensibility	2.07	.047	.504	.243	.744	-.023	1.49
Accentedness	2.31	.028*	.490	.212	.830	.050	1.58
Fluency	3.03	.005*	.843	.279	1.088	.264	1.88
Pronunciation	3.96	.001*	.973	.246	1.422	.529	2.28
Oral Proficiency	3.49	.002*	1.089	.312	1.256	.399	2.08

Note. *$p > .01$ **Cohen's *d* ***CI for the effect size

Table 5 *Kendall's tau-b correlations between temporal variables and listener ratings*

Construct	Articulation rate	*p*	95% CI Lower	Upper
Overall proficiency	1.00*	.017	1.00	1.00
Fluency (AVG)	.949*	.023	.679	1.12
Fluency (AG)	.932*	.021	.282	.996
Pronunciation (AG)	1.00*	.017	1.00	1.00
Proficiency (AG)	1.00*	.017	1.00	1.00

Note. *Correlation is significant at the 0.05 level (2-tailed).

Acoustic correlates of perceived comprehensibility, accentedness, fluency, pronunciation, and oral proficiency

The second purpose of this study was to examine the relationship between the acoustic measures and the ratings of L2 speakers' perceived comprehensibility, accentedness, fluency, pronunciation, and oral proficiency. For this, five acoustic measures for each audio sample were correlated with their mean average of ratings for each construct across participants using Kendall's tau-b correlation. For the variables investigated, nonsignificant findings were found for phonation-time ratio, MLR, number of pauses/total time, number of pauses/speech time, and mean pausing duration ($p > .05$). The only acoustic variable which was associated with listeners' ratings was articulation rate (Table 5). Although the sample size was small, the critical values were high (Plonsky & Oswald, 2014).

Discussion

The present study reveals two important findings. First, listeners judged L2 speech to be more fluent, more comprehensible, and less accented when they saw the speakers (audio-visual mode) than when they only listened to them (audio mode). However, no significant differences in comprehensibility ratings were noted between the two modes. Speakers' pronunciation and oral proficiency were also rated higher in the former mode. Based on the findings, non-verbal information was found to be helpful (cf. Krauss et al., 1991, p. 752; Thompson & Rubin, 1996, p. 333) and not distracting for listeners (cf. Buck, 2001). The findings indicate that listeners' perceptions of fluency and accentedness are positively affected when they see the L2 speaker being listened to. However, Rubin (1992) pointed out that observing the speakers in addition to listening to them can also result in stereotyping of speakers and lead to less positive evaluations of their accents. On the other hand, the positive contribution of video mode in ratings of spoken text was supported by Nakatsuhara et al. (2021) who indicated how the test takers' use of visual cues and communication strategies compensated for issues in pronunciation and fluency. Tsunemoto et al. (2022) found that visual information positively affected the ratings with lower accentedness ratings and higher comprehensibility ratings while fluency scores were not affected. In the present study, however, fluency ratings were positively affected while com-

prehensibility ratings were not significantly different despite being higher in the video mode. Previous studies have already shown that video provides the raters with a whole picture of speakers' way of communicating (e.g., Salaberry & Kunitz, 2019) by increasing raters' confidence while causing more variability and tolerance in ratings, which may require further investigation (Nakatsuhara et al., 2021). Similarly, unlike some previous studies mentioned, the present study does not look at the contribution of a specific visual cue in ratings under any real testing conditions; however, the findings are an indication of how the use of body language and gestures might contribute to maintaining communication between ITAs and L1 speakers. In this vein, the scoring of high-stakes oral examinations like the TOEFL or the IELTS was already examined by several researchers (e.g., Nakatsuhara et al., 2021), and warrants further investigation to show what is, in fact, being evaluated. By including visual information, scores may increase, but fresh information may not be captured. This also depends on the purpose of speech ratings, which may or may not include an assessment of communication skills and strategies, but as Nakatsuhara et al. (2021) also note, although video mode is preferred, regardless of the mode, standardization is crucial.

The findings also revealed that there was some overlap between acoustic measurements and perceptual ratings, which specifically highlighted the relationship between articulation rate and listener ratings indicated in previous research (Crowther et al., 2018; De Jong, 2018; Kang, 2012; Magne et al., 2019; Saito et al., 2017b). This further demonstrates the speech features that listeners utilize when making assessments regarding speakers' proficiency and fluency (Segalowitz, 2016), with articulation rate - a non-composite fluency index - having the most significant association. However, due to the differences in the selection, annotation, and analysis of fluency measures (see Bosker et al., 2013), the predictive power of temporal measures has been shown to vary and thus should be interpreted with caution (see Suzuki et al., 2021). Although quantitative measures as such might seem difficult to interpret for L2 teachers (Foster, 2020), there are various ways to apply this information in classrooms. Some of these include explicit instruction through awareness-raising hands-on activities (Rossiter et al., 2010) or using an input approach with repeated reading activities (Gorsuch, 2011). Learners may be made aware of the fact that disfluency and hesitations might cause them to be perceived as "less powerful" and "less convincing" than those with fewer disfluency while pointing out the difference between natural and unnatural hesitations, disfluencies, and pauses (see De Jong, 2018). It should, however, be noted that L2 fluency development might be a slow process because it requires developing automaticity during speech production by putting an increased cognitive load on L2 speakers (Ejzenberg, 2000). Learners and teachers should be conscious of the fact that their time and effort might not yield immediate results, as L2 speech development depends on many factors such as age, increased practice and experience, and motivation (Saito et al., 2017a; Saito et al., 2018).

As an implication for ITA testing, the results further contribute to the debate in the literature regarding the inclusion of untrained raters in ITA testing for which the previous literature has been inconclusive. The present study reveals that undergraduate students can successfully judge ITA speech performances despite being untrained. This supports previous research (Derwing et al., 2004; Isaacs, 2008) although Hsieh (2011) suggested that undergraduate students remain unofficial rather than official raters. Still, whether to include undergraduate students as official raters is a question that needs further exploration, but one that deserves consideration. It should, however, be noted that student ratings may not be as "sensitive" in borderline cases where scores truly matter for remediation, and that might cause problems as the ITA testing in countries where it is legislated by the state.

Conclusion

The current study revealed that among L1 English listeners, ratings of L2 speech were higher for measures of fluency, accentedness, pronunciation, and oral proficiency in audio-visual mode compared to audio mode, while there were no significant differences in ratings for comprehensibility. Furthermore, articulation rate was found to be the only acoustic measure associated with the ratings. However, the findings should be interpreted with caution and need to be confirmed, considering the small number of participants involved and the specific context in which the study was conducted.

Despite the limitations, the present study presents crucial pedagogical implications for ESL professionals and ITA trainers calling for more attention to be paid to oral proficiency and its assessment

modes. Given that visual information plays such a critical role in the perceptual judgments of L2 pronunciation, language practitioners might underscore this aspect in their classrooms to encourage their students to use their body language, gestures, eye movement, and facial expressions as well as the classroom tools more effectively during their oral presentations and lectures. Pedagogical implications might be extended to include the preparation of online/blended materials and resources involving L2 speakers in which visual information is used effectively. Besides, given that articulation rate was found to be related to perceptions of L2 various speech constructs, instruction specifically geared towards fostering L2 speakers' fluency should be encouraged.

Finally, while helping ITAs improve their oral proficiency and teaching skills is important, it is also critical to provide undergraduate students with opportunities to increase their global, cultural, and linguistic awareness and appreciation of intercultural differences through various programs that will promote mutual understanding (see Kang & Moran, 2019). This way, both sides show effort to pave the way for positive learning and teaching experiences to take place in college classrooms.

Acknowledgments

The author wishes to thank the reviewers and the editor for all the valuable and insightful comments and suggestions which contributed to improving the quality of the manuscript.

References

Bejar, I., Douglas, D., Jamieson, J., Nissan, S., & Turner, J. (2000). TOEFL *2000 listening framework: A working paper* (TOEFL Monograph Series- MS19). https://www.ets.org/Media/Research/pdf/RM-00-07-Bejar.pdf

Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer [Computer program]. Version 5.3.47. http://www.praat.org/

Bosker, H. R., Pinget, A. F., Quené, H., Sanders, T., & de Jong, N. H. (2013). What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing*, *30*(2), 159–175. https://doi.org/10.1177/0265532212455394

Buck, G. (2001). *Assessing listening*. Cambridge University Press. https://doi.org/10.1017/CBO9780511732959

Burgess, C. S. (2001). *Speaking rate, fluency, and accentedness in the speech of second language learners* (Unpublished doctoral dissertation). Simon Fraser University, Canada.

Burgoon, J. K., Guerrero, L. K., & Floyd, K. (2016). *Nonverbal communication.* Routledge

Chambers, F. (1997). What do we mean by fluency? *System*, *25*, 535–544. https://doi.org/10.1016/S0346-251X(97)00046-8

Conlan, C. J., Bardsley, W. N., & Martinson, S. H. (1994). *A study of intra-rater reliability of assessments of live versus audio-recorded interviews in the IELTS speaking component.* A report by the International Editing Committee of the IELTS.

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

Dahl, T. I., & Ludvigsen, S. (2014). How I see what you're saying: The role of gestures in native and foreign language listening comprehension. *Modern Language Journal*, *98*(3), 813–833. https://doi.org/10.1111/modl.12124

De Jong, N., Steinel, M., Florijn, A., Schoonen, R., & Hulstijn, J. (2012). Facets of speaking proficiency. *Studies in Second Language Acquisition, 34*(1), 5–34. https://doi.org/10.1017/S0272263111000489

De Jong, N. H. (2018). Fluency in second language testing: Insights from different disciplines. *Language Assessment Quarterly*, *15*(3), 237–254. https://doi.org/10.1080/15434303.2018.1477780

De Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*, *41*(2), 385–390. https://doi.org/10.3758/BRM.41.2.385

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility. *Studies in second language acquisition*, *19*(1), 1–16. https://doi.org/10.1017/S0272263197001010

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly, 39*, 379–397. https://doi.org/10.2307/3588486

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research* (Vol. 42). John Benjamins. https://doi.org/10.1075/lllt.42

Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language Learning*, *54*(4), 655–679. https://doi.org/10.1111/j.1467-9922.2004.00282.x

Drijvers, L., & Özyürek, A. (2017). Visual context enhanced: The joint contribution of iconic gestures and visible speech to degraded speech comprehension. *Journal of Speech, Language, and Hearing Research*, *60*(1), 212–222. https://doi.org/10.1044/2016_JSLHR-H-16-0101

Drijvers, L., & Özyürek, A. (2019). Non-native listeners benefit less from gestures and visible speech than native listeners during degraded speech comprehension. *Language and Speech*, *63(2),* 209 –220. https://doi.org/10.1177/0023830919831311

Educational Testing Service. (2004). IBT (Internet Based Testing)/Next Generation TOEFL, Independent Speaking Rubric. Retrieved March 5th, 2019, from http://www.ets.org/Media/Tests/TOEFL/pdf/Speaking_Rubrics.pdf

Ejzenberg, R. (2000). The juggling act of oral fluency: A psycho-sociolinguistic metaphor. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 287–313). The University of Michigan Press.

Foster, P. (2020). Oral fluency in a second language: A research agenda for the next ten years. *Language Teaching*, *53*(4), 446–461. https://doi.org/10.1017/s026144482000018x

Fukumura, K., van Gompel, R. P., & Pickering, M. J. (2010). The use of visual context during the production of referring expressions. *The Quarterly Journal of Experimental Psychology*, *63*(9), 1700–1715. https://doi.org/10.1080/17470210903490969

Gass, S., & Varonis, E. M. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning*, *34*, 65–89. https://doi.org/10.1111/j.1467-1770.1984.tb00996.x

Gorsuch, G. (2016). International teaching assistants at universities: A research agenda. *Language Teaching*, *49*(2), 275–290. https://doi.org/10.1017/S0261444815000452

Gorsuch, G. J. (2011). Improving speaking fluency for international teaching assistants by increasing input. *TESL-EJ*, *14*(4).

Harris, T. (2003). Listening with your eyes: The importance of speech-related gestures in the language classroom. *Foreign Language Annals*, *36*(2), 180–187. https://doi.org/10.1111/j.1944-9720.2003.tb01468.x

Hasan, A. S. (2000). Learners' perceptions of listening comprehension problems. *Language Culture and Curriculum*, *13*(2), 137–153. https://doi.org/10.1080/07908310008666595

Hsieh, C. N. (2011). Rater effects in ITA testing: ESL teachers' versus American undergraduates' judgments of accentedness, comprehensibility, and oral proficiency. *Spaan Fellow Working Papers*, *47*.

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. https://doi.org/10.1111/lang.12451

Isaacs, T. (2008). Towards defining a valid assessment criterion of pronunciation proficiency in non-native English-speaking graduate students. *The Canadian Modern Language Review*, *64* (4), 555–580. https://doi.org/10.3138/cmlr.64.4.555

Isaacs, T., & Trofimovich, P. (2011). Phonological memory, attention control, and musical ability: Effects of individual differences on rater judgments of L2 speech. *Applied Psycholinguistics*, *32*, 113–140. https://doi.org/10.1017/S0142716410000317

Jenkins, S., & Parra, I. (2003). Multiple layers of meaning in an oral proficiency test: The complementary roles of nonverbal, paralinguistic, and verbal behaviors in assessment decisions. *Modern Language Journal 87*(1), 901–107. https://doi.org/10.1111/1540-4781.00180

Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, *38*(2), 301–315. https://doi.org/10.1016/j.system.2010.01.005

Kang, O. (2012). Impact of rater characteristics and prosodic features of speaker accentedness on ratings of international teaching assistants' oral performance. *Language Assessment Quarterly, 9*(3), 249–269. https://doi.org/10.1080/15434303.2011.642631

Kang, O., & Moran, M. (2019). Enhancing communication between ITAs and US undergraduate students. In S. Looney & S. Bhalla (Eds.), *A transdisciplinary approach to international teaching assistants* (pp. 82–99). Multilingual Matters. https://doi.org/10.21832/9781788925556-006

Kang, O., & Rubin, D. (2009). Reverse linguistic stereotyping: Measuring the effect of listener expectations on speech evaluation. *Journal of Language and Social Psychology, 28*(4), 441–456. https://doi.org/10.1177/0261927X09341950

Kang, O., Rubin, D., & Lindemann, S. (2015). Mitigating US undergraduates' attitudes toward international teaching assistants. *TESOL Quarterly*, *49*(4), 681–706. https://doi.org/10.1002/tesq.192

Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal*, *94*(4), 554–566. https://doi.org/10.1111/j.1540-4781.2010.01091.x

Kennedy, S., & Trofimovich, P. (2008). Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context. *Canadian Modern Language Review/La Revue canadienne des langues vivantes*, *64*(3), 459–489. https://doi.org/10.3138/cmlr.64.3.459

Kormos, J. (2006). *Speech production and second language acquisition*. Lawrence Erlbaum.

Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, *32*(2), 145–164. https://doi.org/10.1016/j.system.2004.01.001

Krauss, R., Morrel-Samuels, I., & Colasante, C. (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology*, *61*(5), 743–754. https://psycnet.apa.org/doi/10.1037/0022-3514.61.5.743

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, *40*, 387–417. https://doi.org/10.1111/j.1467-1770.1990.tb00669.x

Levis, J. M. (2006). Pronunciation and the assessment of spoken language. In R. Hughes (Ed.), *Spoken English, TESOL and applied linguistics: Challenges for theory and practice* (pp. 245–270). Palgrave. https://doi.org/10.1057/9780230584587_11

Lindemann, S. (2003). Koreans, Chinese or Indians? Attitudes and ideologies about non-native English speakers in the United States. *Journal of Sociolinguistics*, *7*(3), 348–364. https://doi.org/10.1111/1467-9481.00228

Lindemann, S., & Subtirelu, N. (2013). Reliably biased: The role of listener expectation in the perception of second language speech and its implications for research and pedagogy. *Language Learning, 63*, 567–594. https://doi.org/10.1111/lang.12014

Magne, V., Suzuki, S., Suzukida, Y., Ilkan, M., Tran, M., & Saito, K. (2019). Exploring the dynamic nature of second language listeners' perceived fluency: A mixed-methods approach. *TESOL Quarterly, 53*(4), 1139–1150. https://doi.org/10.1002/tesq.528

Meyer, K. R., & Mao, Y. (2014). Comparing student perceptions of the classroom climate created by US American and international teaching assistants. *Higher Learning Research Communications*, *4*(3), 12–22. https://doi.org/10.18870/hlrc.v4i3.206

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. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language learning*, *49*, 285–310. https://doi.org/10.1111/0023-8333.49.s1.8

Munro, M. J., & Derwing, T. M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech: The role of speaking rate. *Studies in Second Language Acquisition*, *23*(4), 451–468. https://doi.org/10.1017/s0272263101004016

Nakatsuhara, F., Inoue, C., & Taylor, L. (2021). Comparing rating modes: Analysing live, audio, and video ratings of IELTS speaking test performances. *Language Assessment Quarterly*, *18*(2), 83–106. https://doi.org/10.1080/15434303.2020.1799222

Pinget, A.-F., Bosker, H. R., Quené, H., & de Jong, N. H. (2014). Native speakers' perceptions of fluency and accent in L2 speech. *Language Testing*, *31*(3), 349–365. https://doi.org/10.1177/0265532214526177

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning*, *64*, 878–912. https://doi.org/10.1111/lang.12079

Rossiter, M. J. (2009). Perceptions of L2 fluency by native and non-native speakers of English. *Canadian Modern Language Review/La Revue Canadienne Des Langues Vivantes*, *65*(3), 395–412. https://doi.org/10.3138/cmlr.65.3.395

Rossiter, M. J., Derwing, T. M., Manimtim, L. G., & Thomson, R. I. (2010). Oral fluency: The neglected component in the communicative language classroom. *Canadian Modern Language Review*, *66*(4), 583–606. https://doi.org/10.1353/cml.2010.0010

Rost, M. (2011). *Teaching and researching listening* (2nd ed.). Pearson.

Rubin, D. L. (1992). Nonlanguage factors affecting undergraduate's judgments of nonnative English-speaking teaching assistants. *Research in Higher Education*, *33*(4), 511–531. https://doi.org/10.1007/BF00973770

Rubin, J. (1995). The contribution of video to the development of competence in listening. In D. Mendelsohn, & J. Rubin (Eds.), *A guide for the teaching of second language listening* (pp. 151–165). Dominie Press.

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., Dewaele, J.-M., & Hanzawa, K. (2017a). A longitudinal investigation of the relationship between motivation and late second language speech learning in classroom settings. *Language and Speech*, *60*(4), 614–632. https://doi.org/10.1177/0023830916687793

Saito, K., Ilkan, M., Magne, V., Tran, M. N., & Suzuki, S. (2018). Acoustic characteristics and learner profiles of low-, mid- and high-level second language fluency. *Applied Psycholinguistics*, *39*(3), 593–617. https://doi.org/10.1017/S0142716417000571

Saito, K., Trofimovich, P., & Isaacs, T. (2017b). Using listener judgments to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics*, *38*(4), 439–462. https://doi.org/10.1093/applin/amv047

Saito, K., Webb, S., Trofimovich, P., & Isaacs, T. (2016). Lexical correlates of comprehensibility versus accentedness in second language speech. *Bilingualism: Language and Cognition*, *19*(3), 597–609. https://doi.org/10.1017/S1366728915000255

Salaberry, M. R., & Kunitz, S. (Eds.). (2019). *Teaching and testing interactional competence: Bridging theory and practice.* Routledge.

Sanchez, C. A., & Khan, S. (2016). Instructor accents in online education and their effect on learning and attitudes. *Journal of Computer Assisted Learning*, *32*(5), 494–502. https://doi.org/10.1111/jcal.12149

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, *14*, 357–385. https://doi.org/10.1017/S0272263100011189

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. Routledge.

Segalowitz, N. (2016). Second language fluency and its underlying cognitive and social determinants. *International Review of Applied Linguistics in Language Teaching*, *54*(2), 79–95. https://doi.org/10.1515/iral-2016-9991

Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, *55*(4), 661–699. https://doi.org/10.1111/j.0023-8333.2005.00320.x

Suzuki, S., Kormos, J., & Uchihara, T. (2021). The relationship between utterance and perceived fluency: A meta-analysis of correlational studies. *The Modern Language Journal*. https://doi.org/10.1111/modl.12706

Tavakoli, P. (2016). Fluency in monologic and dialogic task performance: Challenges in defining and measuring L2 fluency. *International Review of Applied Linguistics in Language Teaching*, *54*(2), 133–150. https://doi.org/10.1515/iral-2016-9994

Tavakoli, P., Nakatsuhara, F., & Hunter, A. M. (2020). Aspects of fluency across assessed levels of speaking proficiency. *The Modern Language Journal*, *104*(1), 169–191. https://doi.org/10.1111/modl.12620

Tavakoli, P., & Wright, C. (2020). *Second language speech fluency: From research to practice*. Cambridge University Press.

Thompson, I., & Rubin, J. (1996). Can strategy instruction improve listening comprehension? *Foreign Language Annals*, *29*(3), 331–342. https://doi.org/10.1111/j.1944-9720.1996.tb01246.x

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

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

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. https://doi.org/10.1111/modl.12779

Wood, D. (2001). In search of fluency: What is it and how can we teach it? *Canadian Modern Language Review, 57*, 573–589. https://doi.org/10.3138/cmlr.57.4.573

[1] However, Plakans' (1997) study showed that American college students feel more emphatic towards ITAs as students come close to graduation.

[2] IELTS Speaking band descriptors (public version) is available at https://www.ielts.org/-/media/pdfs/speaking-band-descriptors.ashx?la=en