

A MIXED-METHODS STUDY EXPLORING PERCEPTIONS OF SPEECH FLUENCY

By Kent Williams, Carleton University

Enhancing English language learners' speech fluency is often a key learning outcome in communicative language classrooms. Notably, how fluent a learner's speech is has been shown to affect how comprehensible it is (Derwing, Rossiter, Munro, & Thomson, 2004). For this reason, it is not surprising that fluency has long been an integral component of both high-stakes and low-stakes oral proficiency assessment rubrics (Fulcher, 2003). Decisions that are made based on the results of these assessments may have real-world implications on test-takers' lives. Thus, it is important to understand which features of speech influence how fluency is perceived in order to enhance the validity of fluency assessments. In this study, although the participants reported that a wide range of temporal, non-temporal, and even non-linguistic features of speech influenced how they perceive fluency, it would seem as though, the speed of speech and the percentage of time speaking most strongly influenced how they assessed it.

But what exactly is fluency? According to Lennon (1990), fluency is referred to in both broad and narrow terms. In the broadest sense of the term, fluency equates to overall language proficiency as in "I can speak three languages fluently!" However, in the realm of second language instruction and assessment, fluency is often defined much more narrowly as the overall speed and flow of speech. Yet, even within this narrow realm, definitions seem to vary widely, which can be problematic for assessors. Chambers (1997) highlights this problem by stating that "it cannot be assumed that we all share the same definition of fluency. Otherwise the validity of the judgements made by assessors is seriously in question. (p. 543)"

Much research on fluency has involved investigating temporal variables of speech in terms of speed, pauses, and repairs. Since the 1970s, second language researchers have examined a wide variety of temporal measures including speech rate (number of syllables/duration, including silent pauses), articulation rate (number of syllables/duration, excluding silent pauses), mean length of runs (average number of syllables/utterance), the number, length, and location of silent and filled pauses, and the number and type of repairs. In the early 1990s, Lennon (1990), Riggenbach (1991), and Freed (1995) began correlating these temporal measures with overall impressions of fluency, as assessed on Likert scales. Their results revealed that, depending on the context, certain temporal measures of fluency seem to exert a degree of influence over raters' judgements of speech fluency. On the whole, these results indicate that increased speed and less hesitancy may lead to higher fluency

ratings, but the relationship is not necessarily linear (Fulcher, 2003). Moreover, fluency judgements may be complicated by a number of contextual factors including individual speech style, the speech task, the speaker's willingness to communicate, and the speaker's familiarity with the topic, situation, and conversation partner.

So what features of speech do assessors attend to when they make overall judgements about learners' ability to speak fluently? Although much research has examined how quantitative measures of speed, pauses, and repairs affect fluency judgments (See Prefontaine, Kormos, & Johnson 2016, for a review of research), not much research (Brown, 2007 is a notable exception) has examined how raters, in their own words, perceive and assess fluency. Therefore, this study set out to use both quantitative (e.g. temporal variables) and qualitative (e.g. interviews) methods to examine this area further.

It is also not well-known how English as a Second Language (ESL) learners perceive fluency. If fluency attainment is a desired outcome of language instruction, then ESL learners should have a fuller understanding of what fluency is, what it is comprised of, and how to go about attaining it. With some exceptions (Prefontaine, 2013, for example), so far, there has been little research into how ESL learners perceive fluency.

To help to answer these questions, this research sought to understand which features of speech influence both assessors' and learners' perceptions of speech fluency. More specifically, the research aimed to answer two research questions: (1) what are perceptually salient features of speech fluency, according to expert raters and intermediate- to advanced-level ESL learners enrolled in a Canadian university; and (2) are these features reflected in temporal measures of speech?

Method

This study incorporated a two-phase, mixed-methods, convergent parallel design (Creswell, 2009). Through this design, both quantitative and qualitative data were collected and analysed simultaneously. Then, the results were merged and interpreted.

All research took place at a medium-sized Canadian university. After the university's research ethics board provided clearance for research to begin, volunteers were invited to participate in the study through emails, in-class recruitment speeches, and in-person. There were two groups of participants: (1) ESL learners ($n = 6$) and expert raters ($n = 2$). The ESL learner-participants were enrolled in regular university programs at the undergraduate and graduate level. The rater-participant group had over 15 years of rating experience and over 25 years of ESL teaching, teacher-training, and testing experience.

In phase one, interviews and test simulations of task one of the Oral Language Test (OLT), which is the speaking component of the Canadian Academic English Language (CAEL) assessment, were conducted individually with six ESL speakers. The participants were first asked these pre-test questions: (1) What does speech fluency mean to you? (2) In your

opinion, what is the difference between beginner, intermediate, and advanced levels of fluency?

Then, participants performed the OLT test simulation, which required them to speak for one minute about their previous English language learning experiences. Afterwards, the participants listened to their speeches and rated their performances according to the OLT rubrics (Paragon Testing Enterprises, 2015) and the Common European Framework of Reference (CEFR) rubrics for Fluency (Council of Europe, 2001). The OLT rubrics consisted of a nine-point Likert scale (10–90), which included corresponding descriptors for each level. The descriptors included references to fluency alongside other features of oral proficiency. The CEFR rubrics consisted of performance benchmarks, which were coded by the researcher to a six-point scale (1–6). These descriptors referred specifically to the temporal features of fluency. These rubrics were translated for intermediate-level participants in order to ensure that the rubrics could be understood more accurately. These translations were conducted by associates of the researcher and they were double-checked by a second reviewer. Once the participants assessed their performances according to the OLT rubrics, they were asked these post-test questions: (1) What do you notice about your speech, in terms of fluency? (2) Why did you assign that score to your speech?

In phase two, interviews were conducted with two expert raters who also evaluated the speeches. The raters were asked the same pre-test questions as the learners. Then, the raters were asked to evaluate the learners' speeches using the OLT rubrics. The raters were then asked the same post-test questions as the learners.

The responses from both participant groups were transcribed and then coded according to techniques proposed by Saldaña (2009). In the first cycle coding procedure (in-vivo coding), relevant quotations were extracted from the transcripts. The second-cycle coding technique (pattern coding) involved identifying connections between these quotations, resulting in the development of themes, categories, and sub-categories.

Results and Discussion

Question 1: According to the participants, what are perceptually salient features of speech fluency?

Individually, salient features varied across participants, meaning that some features of speech were more indicative of speech fluency than others. For instance, some participants mentioned that learners' level of fluency was highly related to their range and use of vocabulary and grammar, whereas other participants noted that one's fluency level depended primarily on how comfortable and familiar a person is with the topic, situation, or conversation partner.

Collectively, perceptions can be situated somewhere between Lennon's (1990) broad and narrow senses of fluency. In other words, overall, participants did not equate fluency with

oral proficiency, but they also did not perceive fluency as being comprised solely of its temporal features. These results suggest that the narrowest definition of fluency as “the speed and flow of speech” does not fully represent the fluency construct.

Categorically, participants made inferences about how fluent the speaker is by the perceived degree of: a) automaticity (how efficiently the learner is able to access linguistic resources, such as vocabulary and grammatical structures); b) comfort in one’s ability to speak English; c) grammatical competency (the range of available linguistic resources); d) speed and flow; e) contextual and cultural familiarity with the topic, situation, or conversational partner; and f) receptivity of speech (how well the message is received by the listener).

Question 2: Are these features reflected in temporal measures of speech?

Comparing raters’ and test-takers’ assessments on the OLT scale with temporal measures of speech required the following: (1) examining the interrater reliability of their assessments ($r = .92$); (2) calculating the temporal measures of individual speeches and their corresponding mean ratings; and (3) correlating the temporal measures with the mean ratings.

The following temporal measures were chosen for this study because they have been used repeatedly in a number of previous studies: Speech Rate (SR); Articulation Rate (AR); Phonation-Time Ratio (PTR) (percentage of time speaking); Mean Length of Runs (MLR); Number of Silent Pauses (SP); and the Number of Filled Pauses (FP).

The results indicated that SR ($r = .86, p < .027$) and PTR ($.85, p < .03$) were most strongly, and significantly correlated with raters’ assessments on the OLT and the CEFR. Additionally, on both scales, AR ($r = .79$), MLR ($r = .79$), and SP ($r = -.71$) also produced moderately-high correlations. The number of filled pauses was not strongly correlated ($r = .52$). These results seem to suggest that the speed of speech (SR), and the percentage of time speaking (PTR) influence how raters assess overall proficiency (OLT) and temporal fluency (CEFR).

As for the test-takers’ assessments, PTR correlated highly with test-takers’ assessments on the OLT ($r = .85, p < .03$), and SR correlated highly with test-takers assessments on the CEFR ($r = .86, p < .02$). Moderate to low correlations were discovered for all other temporal measures. Overall, comparing the results from both participant groups, it would seem as though the speed of speech and the percentage of time speaking most strongly influenced participants’ assessments on this task.

To review, the qualitative results to the first research question suggest the importance of non-temporal and even non-linguistic (e.g. the participants’ perceived level of comfort) features of speech on affecting fluency judgements. The results relating to the second research question suggest the importance of temporal features such as speech rate on affecting fluency judgments. Merging these two results highlights how both temporal and non-temporal features of speech may be interrelated. Table 1 highlights these temporal and non-temporal connections as expressed by the participants in this study.

Table 1: Summary of findings by participant type

Themes	Temporal/Non-temporal Connections	Participant
Automaticity	“Her pace is very slow and ponderous and you can feel her thinking through and trying to speak grammatically.”	Rater
	“When I speak some English is very slow. I try not to make any mistake.”	Learner
Comfort	“There’s a great deal of comfort just speaking. There’s no pausing, halting.”	Rater
	“Fluent but not as comfortable [as another speaker]. It doesn’t flow quite as smoothly.”	Rater
Grammatical competency	“Vocabulary range is high, um, you know, cause there’s no hesitancy.”	Rater
	“It’s all packaged units, not that she can’t communicate effectively but it doesn’t have the same kind of flow.”	Rater
Contextual/cultural familiarity	“But if there’s a difficult subject or unfamiliar situation, it can hinder my smooth flow.”	Learner
	“When it’s certain subject not familiar with me I’m really overwhelmed and I can’t control myself. So I answer sometimes too much pause and hesitation.”	Learner
Receptivity	“At the advanced level, you’re able to relax as an interlocutor because you’re actually communicating. You forget that you’re in a testing situation because it’s (the learner’s speech) is just floating. So if I had to take one big holistic charge (about categorizing fluency), I would say it’s flow.”	Rater
	“When assessing fluency, there’s a point in the process where you have to think about what you’re doing, how comfortable you are as the interlocutor.”	Rater

Pedagogical implications

The purpose of the study was to examine how speech fluency is perceived by expert raters and ESL learners enrolled in a Canadian university. The results indicate that temporal measures and non-temporal measures appear to be inherently interrelated, further revealing the complexity of the speech fluency construct. Therefore, as fluency affects speech comprehensibility, and as it continues to be an integral component of oral proficiency testing rubrics both within and beyond the classroom, it is important to understand which features of speech influence how ESL practitioners define, categorize, analyse, and evaluate speech fluency. It would therefore be worthwhile for instructors to reflect on how they perceive and assess fluency as this reflection could help to increase their confidence in making valid judgments about fluent ability. Additionally, it would be equally worthwhile to elicit how students perceive fluency in order to raise their awareness of how fluency is defined, categorized, analysed, and evaluated. As Chambers (1997) notes, it is not helpful to simply ask students to speak faster and pause less. If both instructors and students become more aware of what fluency is, and what it is comprised of, it is quite possible that instructors and students may have a greater understanding of how to attain it.

References

- Brown, A. (2007). An investigation of the rating process in the IELTS oral interview. In L. Taylor & P. Falvey (Eds.), *IELTS collected papers: Research in speaking and writing assessment*. (pp. 98-141). Cambridge: Cambridge University Press.
- Chambers, F. (1997). What do we mean by fluency? *System*, 25(4), 535-544.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, UK: Cambridge University Press.
- Creswell, J. (2009). *Research Design: Qualitative, quantitative, and mixed methods approaches, 3rd Edition*. Thousand Oaks, CA: Sage.
- Derwing, T., Rossiter, M., Munro, M., & Thomson, R. (2004). Second language fluency: Judgments on different tasks. *Language Learning*, 54(4), 655-679.
- Freed, B. (1995). What makes us think that students who study abroad become fluent? In B. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 123-48). Amsterdam: John Benjamins.
- Fulcher, G. (2003). *Testing second language speaking*. London: Longman.
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40(3), 387-417.
- Paragon Testing Enterprises (2015). *Oral Language Sample Test*. Retrieved from: https://www.cael.ca/wcontent/uploads/2015/10/OLT_practice_test_October_2002.pdf
- Prefontaine, Y. (2013). Perceptions of French fluency in second language speech production. *The Canadian Modern Language Review*, 69(3), 324-348.
- Prefontaine, Y., Kormos, J., & Johnson, D. E. (2016). How do utterance measures predict raters' perceptions of fluency in French as a second language? *Language Testing*, 33(1), 53-73.
- Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes*, 14(4), 423-441.
- Saldaña, J. (2009). *The coding manual for qualitative researchers*. Los Angeles: Sage.

Author Bio



Kent Williams is a PhD candidate in Applied Linguistics and Discourse Studies at Carleton University and he is also a Sessional Lecturer in the Culture and Language Studies Department at Renison University College at the University of Waterloo. His primary research interests are second language speech proficiency, language testing and assessment, and curriculum and materials development. He has taught English in Canada, South Korea, Vietnam, and the Czech Republic.