Temporal measurements for Second Language speaking fluency

Essa Qurbi [a]

[a] *Najran University, Saudi Arabia*

Received December 18, 2024; Revised December 22, 2024; Accepted December 28, 2024

Abstract. This study investigates the relationship between temporal measurements and second language (L2) speaking fluency, focusing on speech rate, mean length of utterance, and pausing patterns. The study recruited 60 advanced English L2 adult learners (male and female) who were presented with a picture book where each page included only one picture representing a specific sequence of the story's development. The participants' main task was to orally describe the story where their oral descriptions were electronically recorded using PsychoPy software. The primary goal was to determine which of the above-mentioned temporal measurements significantly influence L2 speaking assessment. These participants' spoken samples were evaluated by five English language instructors using standardized L2 speaking assessment rubric. A statistical analysis of was conducted to examine the effect of these three temporal measurements on L2 speaking assessment. The results showed that average words per minute and average pause duration per minute, but not mean length of utterance, were significant factors of L2 speaking fluency assessment. However, the interaction between these factors was not significant. That is, average words per minute and average pause duration per minute factors provide a listener with a positive impression about L2 speakers' speaking performance. These results are consistent with other studies that examined L2 speaking fluency. The findings also underline the importance of listeners' impression when it comes to L2 speaking assessment. In addition, the study highlights the effect of temporal measurements on L2 speaking assessments and provides insights for both evaluators and learners on the factors affecting L2 fluency assessment.

Keywords*: L2 fluency, temporal measurements, speech rate, pausing patterns, mean length of utterance, L2 speaking assessment.*

Курбі Есса. Часові параметри оцінки плавності говоріння другою мовою.

Анотація. Це дослідження вивчає зв'язок між часовими параметрами та плавністю говоріння другою мовою, зосереджуючись на темпі мовлення, середній тривалості висловлювання та середній тривалості пауз за хвилину. У дослідженні взяло участь 60 дорослих чоловіків і жінок, які вивчають англійську мову як другу. Їм було запропоновано книжку з картинками, де на кожній сторінці містилася лише одна картинка, що представляла певну послідовність розвитку сюжету. Завданням учасників було створити історію в усній формі, а їхні усні описи записувалися в електронному

Essa Qurbi, 0000-0002-7073-0602, eabatel@nu.edu.sa

East European Journal of Psycholinguistics, 11(2), 142–153. https://doi.org/10.29038/eejpl.2024.11.2.qur

вигляді за допомогою програмного забезпечення PsychoPy. Головна мета полягала в тому, щоб визначити, які з вищезгаданих часових параметрів суттєво впливають на оцінку усного мовлення на рівні L2. Мовленнєві зразки учасників оцінювали п'ять викладачів англійської мови, використовуючи стандартизовану шкалу оцінювання говоріння на рівні L2. Для вивчення впливу цих трьох часових параметрів на оцінювання говоріння на рівні L2 було проведено статистичний аналіз. Результати показали, що темп мовлення та середня тривалість пауз за хвилину, але не середня довжина висловлювання, були значущими чинниками оцінки вільності говоріння на рівні L2. Однак взаємодія між цими факторами не була значущою. Тобто, фактори середньої кількості слів за хвилину та середня тривалість пауз створюють у слухача позитивне враження про мовлення носіїв другої мови. Ці результати узгоджуються з іншими дослідженнями щодо вільного говоріння другою мовою. Результати також підкреслюють важливість урахування враження слухачів, коли йдеться про оцінку мовлення другою мовою. Крім того, дослідження висвітлює вплив часових параметрів на оцінювання говоріння другою мовою та надає інформацію і тим, хто оцінює, і тим, хто навчається, про фактори, що впливають на оцінювання вільного володіння другою мовою.

Ключові слова: вільне володіння другою мовою, часові параметри, темп мовлення, паузи, середня довжина висловлювання, оцінювання говоріння другою мовою.

Introduction

This empirical study investigated the relationship between some temporal measurements and second-language speaking fluency. Temporal measurements are essential for assessing L2 speaking fluency and offer valuable insights into the speed, smoothness, and timing of speech production. Key measurements such as speech rate, articulation rate, duration of speech segments, pausing patterns and others provide a detailed understanding of second language (L2) speaking fluency. Empirical studies highlight the dynamic nature of L2 fluency development, which is influenced by various factors. For example, De Jong et al (2013) found that linguistic knowledge (e.g., grammar and vocabulary knowledge) and processing speech (e.g., speed of lexical retrieval and speech rate) were strong predictors of L2 fluency of their Dutch L2 speakers when performing L2 speaking tasks. However, the study found that the mean syllable duration was the strongest predictor of L2 speaking fluency. Therefore, the following section provides a literature review of some of the studies that investigated the temporal measurements and their relationship with L2 speaking fluency.

Literature Review

Speaking Fluency

Speaking fluency is an important indicator of second language proficiency. It is often associated with the ease, speed, and smoothness of speech production.

Despite several definitions of L2 speaking fluency, there is no agreed-upon definition by psycholinguists. However, most of these definitions have common features, including fluidity and the speaker's ease of expressing themselves in the second language (e.g., Fillmore, 1979; Freed, 2000; Kormos, 2006). For example, Nation (2014) defined L2 speaking fluency as "the ability to process language receptively and productively at a reasonable speed." (p. 11). It is distinct from overall language proficiency regarding the flow and temporal measurements of speech (Segalowitz, 2010). Rehbein (1987) defined L2 speaking fluency as "the activities of planning and uttering [that are] executed nearly simultaneously by the speaker of the language" (p. 104). This definition considers the spontaneous process of planning and executing speaking.

Temporal Measurements

Temporal measurements for speaking fluency refer to a number of methods used to calculate and analyze the timing aspects of L2 speech. These measurements are indicators of L2 speaking fluency (e.g., Iwashita et al., 2008; Lennon, 1990; Towell et al., 1996). A number of studies have investigated the relationship between temporal aspects and L2 speaking performance, where some of these temporal measurements correlate with L2 speaking fluency (Berto & Galaverna, 2016; García-Amaya, 2009). These measurements can be used individually or in combination to provide a comprehensive evaluation of speaking fluency. Among these temporal measurements are those presented by Segalowitz (2010):

- **Speech Rate (SR)** is the number of syllables or words produced per minute. It reflects the overall speed of speech production. Lennon (1990) identified speech rate as a robust indicator of fluency, finding that faster speech rates correlate with higher fluency ratings. Munro and Derwing (1998) demonstrated that speech rate significantly impacts listeners' evaluations of fluency, with faster rates often perceived as more fluent.
- **Articulation Rate (AR)** measures the number of syllables or words spoken per minute, excluding pauses. This metric provides a more accurate reflection of speech production speed. Derwing et al. (2004) highlighted the articulation rate as a significant predictor of perceived fluency. Cucchiarini, Strik, and Boves (2000) found that the articulation rate is closely related to overall speech intelligibility, indicating its importance in fluency assessment.
- **Mean Length of Utterance (MLU)** calculates the average length of spoken units, typically in words or morphemes. Foster and Skehan (1996)

demonstrated that MLU is a useful metric for assessing L2 complexity and fluency. Ortega (2009) showed that a higher MLU correlates with greater linguistic complexity in L2 speech, indicating that more fluent speakers produce longer and more complex utterances.

- **Phonation Time Ratio (PTR)** is the proportion of time spent speaking compared to the total time of the speech sample. Riggenbach (1991) found the phonation time ratio to be a key indicator of fluency, as fluent speakers tend to have higher ratios of speaking time to total time. Kormos and Dénes (2004) confirmed its importance in distinguishing fluent from non-fluent speakers, showing that more fluent speakers have higher phonation time ratios.
- **Pausing Patterns (PP)** include the frequency, duration, and distribution of speech pauses. Kormos and Dénes (2004) showed that excessive pausing correlates with lower fluency ratings. Tavakoli and Skehan (2005) found that strategic planning can reduce the frequency and length of pauses in L2 speech, suggesting that more fluent speakers can manage their pauses more effectively.
- **Turn-Taking Latency (TTL)** measures the time taken to respond in a conversation. Rossiter (2009) found that shorter turn-taking latency is associated with higher fluency, as fluent speakers can respond more quickly in conversational exchanges. Stivers et al. (2009) highlighted cross-linguistic variations in turn-taking latency, emphasizing its role in conversational fluency.
- **Duration of Speech Segments (DSS)** is the length of continuous speech segments between pauses. Longer segments can indicate greater fluency.

Approaches for collecting spoken data include structured interviews, picture description tasks, and narrative retellings. Foster and Skehan (1996) employed narrative retellings to capture spontaneous speech, while Skehan and Foster (1999) used structured tasks to control for complexity and accuracy. These methods provide reliable data for analyzing temporal measurements of fluency.

Empirical research has extensively investigated temporal measurements of L2 fluency, employing various methodologies to examine how these metrics evolve and differ among L2 learners. (e.g., Rossiter, 2009; Tavakoli & Skehan, 2005). Some studies have tracked changes in temporal fluency metrics over time, providing insights into the developmental trajectory of L2 learners. De Jong et al. (2013) examined how temporal measurements of fluency develop over time, finding improvements in speech rate and phonation time ratio as L2 learners gain more proficiency and exposure to L2.

Suzuki, Kormos, and Uchihara (2021) conducted a meta-analysis to examine how utterance features relate to perceived fluency in second language speakers. It posits that faster speech rate and fewer pauses are strongly linked to higher fluency ratings, while longer pauses, frequent self-corrections, and hesitation markers negatively affect fluency perception. The study suggests that fluency is influenced by both speech speed and smoothness, but excessive emphasis on speed can undermine other language aspects like accuracy.

Other studies have compared different proficiency levels to identify patterns and differences in speaking fluency measurements. For example, Freed, Segalowitz, and Dewey (2004) compared learners at different proficiency levels, showing that more proficient speakers showed faster speech rates and fewer pauses. These findings highlight the progressive nature of fluency development, with advanced learners demonstrating more fluent speech characteristics. In addition, Segalowitz and Freed (2004) explored the impact of experiences during study abroad on L2 fluency. The results showed significant gains in speech rates and rates among learners who participated in L2 immersion programs. These studies show the importance of investigating the temporal measurements of L2 speaking, as they can provide insight into understanding both L2 speaking performance and assessment patterns.

However, only three of the above-listed temporal measurements were used in this study: 1) speech rate (SR), 2) pausing pattern (PP), and 3) Mean Length of Utterance (MLU). Specifically, the SR represents the average number of words a speaker produces per minute. This includes pausing time as well. The PP is related to the average time (in milliseconds) of speech pauses per minute. In contrast, the MLU is the length of continuous speech segments (measured by the number of words) between pauses.

Relevance of the Study

Investigating the influence of temporal measurements is crucial for evaluating L2 classroom speaking. A teacher's impression plays a significant role in L2 speaking assessments. Therefore, there is a need for a better and more accurate understanding of this subjective assessment. In other words, it is important to investigate these temporal measurements because they can provide significant insight into understanding both L2 speaking performance and, more importantly, assessment patterns. Measuring L2 speaking fluency based on temporal measurements of L2 speech also contributes to understanding the cognitive processes and linguistic characteristics of proficient L2 speaking.

Methodology

This study tests the relationship between three temporal measurements: 1) the speech rate (SR) (average number of words per minute), 2) Mean Length of Utterance (MLU) (average number of words between pauses), and 3) the pausing pattern (PP) (the average time of pauses per minute), as a factor to explain L2 speaking fluency represented by the speaking assessment score. It is worth mentioning that this paper utilized an AI tool (ChatGPT) for English grammatical correction to enhance clarity and linguistic accuracy of the content.

Research Questions

This study aimed to determine whether there is a significant effect of the three temporal measurements of speaking identified above on L2 speaking fluency. Two questions are relevant to the scope of the study:

- Question 1: Which of the three temporal measurements under investigation has a significant effect on L2 speaking fluency?
- Question 2: Do any of these temporal measurements have a greater impact on L2 speaking fluency than others?

It is assumed that at least one of the three temporal measurements would significantly impact the speaking fluency assessment.

Participants

The study included 60 L2 English speakers who attended college majoring in English as a second language (ESL) at the time of the study in the English Department at Najran University. Their ages ranged from 21 to 25 years; 30 were male and 30 were female. Furthermore, based on an English L2 proficiency questionnaire, they were all in the advanced level of English (average English proficiency level = 7.6/10). These participants ranged from semester 5 (3rd year) to semester 8 (4th year) at their four-year bachelor degree diploma. At this stage, students study a range of advanced English courses like theoretical and applied linguistics, English literature, translation, and others. In addition, five English L2 teachers were recruited to evaluate the speaking performance of the participants.

The five L2 teachers (age mean = 47) were recruited using Prolific's webpage (https://www.prolific.com/) to evaluate the recorded speaking samples of these participants. All raters were native speakers of English, and based on their

responses on the provided questionnaire, and had been teaching English as an L2 at the postsecondary level for an average of 13.3 years (range = 9–18 years of experience). They were provided with the speaking skill rubric of the New York State Education Department to use for assessment. This rubric comprises four speaking categories (five points for each category): appropriate word choice, grammar errors, fluency, and pronunciation. This ensured that the L2 speech raters used the same assessment scale for every speech sample.

Tasks and Procedures

Each participant was presented with a wordless picture book. Each page included only one picture representing a specific sequence of the story's development; this was the speech elicitation task in this study. The participants were asked to orally describe the story in the book. Their responses were electronically recorded using PsychoPy software (https://www.psychopy.org/).

After reading the study instructions and signing the consent form on the computer screen, each participant was asked to press the space bar to move to the next window to start the study. After each participant had completed the task, their spoken responses were automatically saved on the researcher's webpage in Pavlovia, a site created by the PsychoPy team to conduct studies online (https://pavlovia.org/). These responses were then provided to the L2 raters to assess and score each participant's speaking performance. Besides, the speaking samples were analyzed using Audacity software (https://www.audacityteam.org/) and the Online Forced Aligner tool (https://web.uwm.edu/forced-aligner/) to analyze each speech sample's number of words, pausing time, and the mean length of utterance. These are authentic and free tools used for auditory data analysis.

Results

The data were initially analyzed using three-way ANOVA in R software. The L2 raters provided the assessment score as the dependent variable explained by the three temporal measurements investigated in this study (SR, MLU, and PP). The results showed a significant effect of only the SR and the PP. In contrast, neither MLU nor any of the interactions between these factors showed a significant effect on the assessment score. Thus, the MLU factor was dropped from the statistical model to include only the SR and the PP factors in a two-way ANOVA.

Table 1
Descriptive Results

	Mean	**SD**	**Median**	**Range**	**Skew**	**Kurtosis**
Speech_Rate (SR)	62.18	11.28	63.0	39.0	-.10	-1.28
Mean Length of Utterance (MLU)	5.08	.75	5.20	3.40	-.02	-.42
Pausing_Pattern (PP)	5.62	1.77	5.20	6.99	.32	-1.04
Speaking_Score	12.74	2.54	12.50	10.0	-.07	-1.23

The results showed a significant primary effect of both the SR ($F(1,56) = 19.937$, $p < .0001$ at alpha 0.05, effect size = .26) and the PP ($F(1,56) = 97.6$, $p < .0001$ at alpha .05, effect size = .64). However, the interaction between these two factors was not significant ($F(1,56) = 2.124$, $p = .151$, effect size = .04).

Table 2
Two-Way ANOVA Results

	Df	**Sum Sq**	**Mean Sq**	**F value**	***p-value***
Pausing_Pattern (PP)	1	212.24	212.24	97.610	.0000 ***
Speech_Rate (PR)	1	43.35	43.35	19.937	.0000 ***
Pausing_Pattern (PP) x Speech_Rate (PR)	1	4.62	4.62	2.124	.151
Residuals	56	121.76	2.17		

Discussion

This study assessed the effect of three temporal measurements on L2 speaking assessment. Temporal measurements represent the cognitive processes underlying the speaking performance of L1 and L2 speakers. The study included the SR (represented by the average number of words per minute), the MLU (represented by the average number of words between pauses), and the PP (represented by the time of pause between speech runs) to determine their effect on second language (L2) speaking assessment. Sixty English L2 participants participated in this study to provide a spoken response to a given prompt. In addition, five English language teachers were recruited to evaluate the participants' speaking performance. A statistical analysis of the results was

then conducted, including the assessment score as the dependent variable explained by the three temporal measurements as the independent variables of this study. As shown in the results section above, a significant positive effect was observed with the SR and the PP factors on the speaking score, whereas the MLU showed no effect.

The analysis concluded that some temporal measurements could be a reliable indicator of L2 fluency. The PP and SR both represent the pace of speaking production. They provide an impression of the speaker's confidence on the listener's part. On average, the SR was 62.18 words per minute (including pausing time), ranging from 41 to 80 words per minute. The PP was 5.62 milliseconds per minute, ranging from 2.11 to 9.1 for this group of L2 speakers.

Therefore, for Question 1, the results indicate that two of the three temporal measurements under study are significant in L2 speaking assessment. That is, both the SR rate and the PP played a crucial role when evaluating the L2 speaking samples. However, the results showed that the MLU was not significant for the L2 speaking assessment.

Regarding Question 2, the PP had a greater impact on the speaking score, with an effect size of .64 compared to .26 for the SR. In other words, 64 % of the variability in the speaking assessment score is determined by the PP factor, a significant effect size. On the other hand, about 26 % (a medium effect size) of the variability in the speaking assessment score was determined by the SR factor. Although this is a considerable effect size, it is not comparable to that of the PP.

The results show that raters of L2 speech are influenced by the pausing pattern (PP) of the L2 speaker. This could indicate confidence and competence on the speaker's part in that lower pausing time would lead to better L2 speaking fluency. While the speech rate (SR) affected the speaking performance score, its lower impact could be because speaking fast might lead to mispronunciation or misunderstanding, and it is prone to more incorrect word choice.

The results are consistent with other studies that examined different speaking fluency levels. For example, Lennon (1990) found that SR was a strong indicator of speaking fluency in that faster PR correlated with higher fluency ratings. Munro and Derwing (1998) found that PR significantly impacts listeners' evaluations of speaking fluency, with faster rates often perceived as more fluent. In addition, the findings of the present study are similar to those of Kormos and Dénes (2004), which showed that excessive pausing correlates with lower speaking fluency ratings. Tavakoli and Skehan (2005) also found that strategic planning can reduce the frequency and length of pauses in

L2 speech, suggesting that more fluent speakers can manage their pauses more effectively.

A finding of no significant effect of MLU was also presented by Iwashita et al. (2008), who found that mean length of run (i.e., MLU) yielded no significant associations with L2 proficiency level. This differs from that of Ortega (2009), who found that higher MLU correlated with greater linguistic complexity in L2 speech, where more fluent speakers produce longer and more complex utterances. This study, however, found no effect of MLU on L2 speaking fluency ratings.

Conclusions

This study examined the influence of three temporal measurements on the assessment of L2 speaking fluency. The results showed that both speech rate and pausing patterns significantly impacted L2 speaking fluency, as assessed by the speaking scores provided by the English L2 teachers. Specifically, SR, which measures the number of words produced per minute, and PP, the average duration of pauses, emerged as critical factors in determining fluency. Conversely, the Mean Length of Utterance, which measures the average length of spoken units between pauses, did not significantly affect fluency ratings.

The significant effect of SR on L2 speaking fluency observed in this study supports previous research indicating that faster SR are often associated with higher fluency ratings (Lennon, 1990; Munro & Derwing, 1998). A faster SR can reflect a higher level of language processing efficiency and confidence, as fluent speakers tend to produce speech at a faster pace without frequent interruptions. However, it is crucial to recognize that excessively rapid speech might lead to mispronunciations or misunderstandings, which can detract from overall fluency. Therefore, while a moderate to fast SR is beneficial, it should be balanced with clarity and accuracy. Pausing patterns (PP), on the other hand, were found to have an even more substantial effect on fluency ratings than speech rate. The duration of pauses can significantly influence a listener's perception of speaking fluency. Prolonged pauses can suggest hesitation, lack of confidence, or difficulties in language processing, which negatively affect the impression of fluency. This result is consistent with previous studies (e.g., Kormos & Dénes, 2004) that found longer pauses correlate with lower speaking fluency ratings.

The study's finding that MLU did not significantly affect L2 speaking fluency diverges from some prior research suggesting that longer utterances

correlate with greater linguistic complexity and fluency (e.g., Ortega, 2009). Although longer utterances can indicate complexity, they do not necessarily reflect smoothness or ease of speech production. Therefore, while MLU remains a useful measure for assessing linguistic complexity, it might not directly correlate with perceived fluency in the same way as temporal measurements related to SR speed and PP.

The significance of these results is that they identify a number of temporal measurements that L2 speaking evaluators might use when making judgments about L2 speaking fluency. Future research could explore the interaction between different temporal measurements and their collective impact on L2 speaking fluency. For instance, investigating how SR and PP interact with other factors, such as pronunciation and vocabulary use, could provide a more holistic view of L2 speaking fluency. Additionally, longitudinal studies tracking changes in these temporal measurements over time could offer insights into how fluency develops with increased language exposure and practice.

References

Chambers, F. (1997). What do we mean by fluency? *ELT Journal, 51*(1), 22–32.

De Jong, N. H., Steinel, M. P., Florijn, A. F., Schoonen, R., & Hulstijn, J. H. (2013). Linguistic skills and speaking fluency in a second language. *Applied Psycholinguistics, 34*(5), 893–916. https://doi.org/10.1017/S0142716412000522

Derwing, T. M., 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.1467-9922.2004.00277.x

Fillmore, C. J. (1979). On fluency. In D. Kempler & W. S.-Y. Wang (Eds.), *Individual differences in language ability and language behavior* (pp. 85–102). Academic Press.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18*(3), 299–323. https://doi.org/10.1017/S0272263100015065

Freed, B. F., Segalowitz, N., & Dewey, D. P. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition, 26*(2), 275–301. https://doi.org/10.1017/S0272263104262061

Freed, B. (2000). Is fluency, like beauty, in the eyes (and ears) of the beholder? In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 243–265). University of Michigan Press.

Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. (2008). Assessed levels of second language speaking proficiency: How distinct? *Applied Linguistics, 29*(1), 24–49. https://doi.org/10.1093/applin/amm026

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

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

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. MIT Press.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning, 40*(3), 387–417. https://doi.org/10.1111/j.1467-1770.1990.tb00207.x

Munro, M. J., & Derwing, T. M. (1998). The effects of speaking rate on listener evaluations of native and foreign-accented speech. *Language Learning, 48*(2), 159–182. https://doi.org/10.1111/1467-9922.00043

Nation, I. S. P. (2014). Developing fluency. In T. Muller, J. Adamson, P. S. Brown, & S. Herder (Eds.), *Exploring EFL fluency in Asia* (pp. 11–25). Palgrave Macmillan.

New York State LOTE assessments (n.d.). *Sample Assessment Rubric*. Retrieved March 2, 2024, from https://www.nysed.gov/sites/default/files/programs/world-languages/lotecassess.pdf

Ortega, L. (2009). *Understanding second language acquisition*. Routledge.

Rehbein, J. (1987). On fluency in second language speech. In H. W. Dechert & M. Raupach (Eds.), *Psycholinguistic models of production* (pp. 97–105). Ablex Publishing Corporation.

Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes, 14*(4), 423–441. https://doi.org/10.1080/01638539109544857

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

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

Segalowitz, N., & Freed, B. F. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition, 26*(2), 173–199. https://doi.org/10.1017/S027226310426205X

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning, 49*(1), 93–120. https://doi.org/10.1111/0023-8333.00091

Suzuki, S., Kormos, J., & Uchihara, T. (2021). The relationship between utterance and perceived fluency: A meta-analysis of correlational studies. *The Modern Language Journal*, *105*(2), 435-463.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., … Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences, 106*(26), 10587–10592. https://doi.org/10.1073/pnas.0903616106

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239–277). John Benjamins.

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics, 17*(1), 84–119. https://doi.org/10.1093/applin/17.1.84