Journal of Education and Educational Research
ISSN: 2957-9465 | Vol. 10, No. 2, 2024

A Systematic Review of Speech Fluency Measurement in L2 Monologues and Dialogues

Wen Huang

School of Humanities and Social Sciences, Xi'an Jiaotong-Liverpool University, China
Wen.Huang23@student.xjtlu.edu.cn

Abstract: This study synthesises measures of L2 speech fluency in both monologues and dialogues. About monologue fluency measurement, we provide a synthesis of measurement based on Skehan's (2003) three dimensions (speed, breakdown, and repair) and discrepancies among scholars. As for dialogue fluency, the limited number of studies precludes a conclusion of well-established measurement. Furthermore, the reasons for the paucity of studies in this area are elucidated. The study revealed inconsistencies in the standardisation and methodological definitions employed by scholars in monologue fluency research, resulting in a lack of uniformity in the data collected. Concerning dialogues, there is a notable absence of theoretical and operational frameworks.

Keywords: L2 Monologues and Dialogues; Speech Fluency Measurement; Systematic Review; Skehan's (2003)

1. Introduction

Fluency is a central aspect of second language (L2) acquisition, reflecting not only L2 learners' proficiency but also their ability to communicate effectively and naturally. According to Lennon (1990), from an L2 perspective, fluency can be examined from broad and narrow perspectives. Broadly, fluency is seen as a term to measure overall oral proficiency, and being "fluent" is seen as one of the utmost goals of L2 learning. In L2 research, fluency is often referred to in its narrow definition, which is an important component of oral proficiency, encompassing a range of measures such as speed rate, pause, and repair, which are critical for successful communication.

To better study and examine fluency, Segalowitz (2010) breaks it down into three dimensions: cognitive fluency, utterance fluency, and perceived fluency. Cognitive fluency refers to the fluency of an individual's mental processes in processing linguistic information. It refers to the cognitive mechanisms involved in understanding, producing, and using language. Individuals with high cognitive fluency are able to process linguistic information quickly and accurately, which is usually associated with language proficiency and language learning experience (Foster, 2020). Utterance fluency refers to the fluency demonstrated in actual linguistic output and includes features such as rhythm, rate of speech, pauses, repetition, and self-correction of speech. It measures the naturalness and coherence of a language user's oral expression (Kormos & Dénes, 2004; Segalowitz, 2010). Perceived fluency is listeners' subjective evaluation of a speaker's fluency. This judgment can be based on the listener's perception of speech, intonation, rhythm, and overall expression. These three dimensions of fluency are interrelated but distinct and together form a comprehensive concept of fluency. Cognitive fluency is the basis of utterance fluency, while utterance fluency is the direct source of perceived fluency, making utterance fluency a special bridge to connecting the other two concepts (Suzuki & Kormos, 2022). A review of the definitions presented above reveals that the measurement of L2 speech fluency is primarily concerned with utterance fluency. This is because utterance fluency is the only tangible form of a speaker's L2 oral output.

Two modes of speech products are available to assess speech fluency: monologue and dialogue. A monologue is a continuous discourse performed by a single individual without any interaction. In the context of L2 speech fluency, monologue typically entails learners narrating a topic or story. Commonly employed assessment measures include speech rate, pauses, repetitions, and so forth. Dialogue is defined as the interactive process of two or more individuals engaged in communication (Witton-Davies, 2014). In L2 speech fluency measures, dialogue can be structured or free-form, involving question-and-answer sessions, discussions, or role-plays. In addition to the aforementioned measurement, overlap, turn pause, and others can also be assessed to measure dialogue fluency, as dialogues involve interaction between two parties. Nevertheless, in practice, the majority of scholars have opted for the monologue as the method of choice for the collection of speech samples. Monologues are relatively easy to collect and no extraneous factors might disrupt the process. However, the lack of interactivity means that the samples cannot fully reflect the real-world speaking abilities of L2 learners (Foster & Ohta, 2005). The use of dialogues allows for the observation of learners' language adaptation and communication strategies (De Bot et al., 2007). However, in comparison, the collection of dialogues is more challenging and is influenced by the dialogue partner, which may result in a decrease in the credibility of the analysis results.

1.1. The Present Study

The preceding analysis demonstrates that the monologue task was employed with considerably greater frequency than the dialogue task in the L2 fluency measurement. This results in a more abundant task design for monologue tasks than for dialogue tasks, as well as a greater variety of measures of monologue fluency. However, this also gives rise to the question of whether different scholars employ different methodologies when measuring the same variables. As for the dialogue task, in addition to the monologue measure, some specific measures need to be refined because some unique conversation phenomena, such as overlapping, repetition, etc., may occur in dialogue. The objective of this article is to

address the following research questions using a systematic analysis of previous relevant studies.

1. What methods were employed by scholars in the design of monologue tasks to collect speech samples?
2. What methods were employed by scholars in the design of dialogue tasks to collect speech samples?
3. What quantitative measures are currently employed to assess L2 learners' speech fluency in monologue?
4. What quantitative measures are currently employed to assess L2 learners' speech fluency in dialogue?

By the end of this paper, readers will have gained insights into the complexities of measuring L2 speech fluency and the implications of these measures for language learners and educators. This systematic review seeks to provide a foundation for further exploration and refinement of fluency measurement tools and techniques.

1.2. Methods

The literature search was conducted primarily on the Web of Science database, with the "All Fields" classification option selected and the keyword "L2 speech fluency" entered. A total of 370 papers were identified through the search process, which I filtered by the number of citations, the number of papers published by the author in the relevant field, etc. Additionally, I searched for relevant articles by some famous experts in the field, such as Skehan, Segalowitz, Peltonen, Kormos, etc. by choosing Author by category. This study includes 37 articles after filtering.

2. Measurement of Monologue Fluency

2.1. Monologue Task Design

With regard to research on the development of L2, the majority of scholars have chosen to utilise story retellings based on pictures or given segments as monologue tasks to collect speech samples. However, the design of the monologue task varies across scholars, which may also affect the findings to some extent. In the two-year study of Derwing and Munro (2013), the English as a second language participants recounted the same narrative in English from a set of images. The fact that the three task settings were identical is likely to produce a practice effect. This is because even if there is no significant improvement in L2 language proficiency, training on the same task three times may result in improvement. Moreover, the subjects had the opportunity to ask questions, which would certainly affect the results of the study. In order to reduce the potential effects of task familiarity, Huensch and Tracy (2017) made some adjustments to their monologue quest settings. The data were collected through three different picture-based narrative tasks, each assignment is 1 year apart and students are grouped according to counterbalance. Kakitani and Kormos (2024) employed three pictorial prompts in their monologue tasks and the order of the prompts was counterbalanced. The three prompts were structured in a linear sequence and presented a narrative arc that prompted speakers to express the emotions and motivations of the characters. A more holistic approach was employed in the design of the monologue task, to measure participants' fluency in various circumstances (Leonard & Shea, 2017). The design differed for all three tasks. The first task was of the declarative type, the second was of the expository type, and the third was of the picture narration and specified the tense to be used.

2.2. Monologue Fluency Measurement

As research has progressed, scholars have proposed a number of methods for measuring monologue fluency. Skehan (2003) divided utterance fluency into three parts: speed fluency, referring to the rate of speech; breakdown fluency, relating to silent and filled pauses, and repair fluency, concerning repetition, false start, and repair. The following section will present each measure based on the categorisation system employed by Skehan (2003), with a particular focus on the analysis of data from distinct studies.

2.2.1. Speed Fluency Measurement

In the field of speech fluency, speech speed has been a significant measure of investigation. Hadži and colleagues (2012) were among the early contributors to this discussion by excluding unfilled pauses from the total phonation time in their calculation of speech rate. This approach was adopted by Saito et al. (2018), who further refined the calculation by subtracting only filled pauses from the total phonation time. Subsequently, Suzuki and Kormos (2022) provided a clear definition of articulation rate as the mean number of syllables produced per second, divided by the total phonation time, which is the total duration of speech excluding pauses. In more recent studies, both Wright (2020) and Tsunemoto (2023) have included the total silent pause duration in their calculations. They define speech rate as the mean number of syllables produced per second, divided by the total speech duration, including pauses. Speech rate, therefore, is regarded as a composite fluency measure encompassing both speed and breakdown, while articulation rate is a pure speed measure.

There may be some variation in the standardisation of data by different scholars. Tavakoli et al. (2020) and Sanaei et al. (2015) calculated these rates within per minute. Some scholars prefer to count the average number of words, while others are accustomed to using the number of syllables. While these factors may contribute to greater variability in the data across studies, they are unlikely to significantly impact the overall results.

2.2.2. Breakdown Fluency

In breakdown fluency, pause is the most important metadata to measure, and shows a hesitation as opposed to fluency. People tend to think that fewer pauses mean higher fluency, but the reality is not that simple. Pauses are unavoidable in expression. What is more important is what function it serves when it is used by the speaker, and this functional difference may differentiate the fluent and non-fluent speakers. The functions can be categorised into hesitation, grammatical marking, breathing, and so on (Duez, 1985). Thus, in order to distinguish between different functions of pauses, scholars have derived a series of measures.

Firstly, regarding the duration of the pause. Usually, academics classify the length of the pause as between 0.2s and 0.3s. Considering differences in monologue tasks, definitions may vary across studies. For example, given his focus on read-aloud tasks, Tsunemoto (2023) reduced the pause time to 200ms, he posited that the duration of pauses does not need to align with that of spontaneous speech. Hadži et al. (2012) considered pause as a breakdown lasting at least 0.2 to 0.25 seconds. The most widely used pause standard currently is 0.25 seconds (Bosker et al., 2013; Ginther et al., 2010; Judit

Kormos & Dénes, 2004). Some have relaxed the standard to 0.4 seconds (Derwing et al., 2004; Tavakoli & Skehan, 2005).

Secondly, the categorisation of the pause must be considered. Pauses are commonly distinguished as silent pauses, and filled pauses. Some scholars measure the two separately, others combine them. As opposed to a silent pause, a filled pause refers to a pause in the sound made by the use of an intonation such as uh, um, oh, etc. Filled pauses are also believed to be a common editing marker for repairs (Levelt, 1983). Furthermore, the frequency of filled pauses varies considerably depending on the speaker's individual speech patterns. Therefore, it is less emphasised in the measurement compared to silent pauses.

Thirdly, the frequency of pauses should be addressed. The four main methods of calculating pause frequency that is currently prevalent are as follows: the number of pauses per minute, the number of pauses per 100 words/syllables, and the number of pauses per sentence/unit (Giles, 2014). Scholars have employed disparate methodologies for calculating the number of pauses, according to their research objectives and the specific counting conventions they have adopted. For instance, Möhle (1984) counted the number of pauses per 100 syllables, Freed (1995) enumerated the number of disfluent, unfilled pauses per 100 words, and Kormos and Dénes (2004) recorded the number of silent pauses per minute. Peltonen (2018) calculated the number of syllables divided by the number of runs (a run consists of speech between silent pause). Tsunemoto (2023) narrowed down pauses into silent pauses. A commonly used measure of pause frequency, the mean length of run, can be driven by counting the mean number of syllables/words produced in utterances between pauses (Suzuki & Kormos, 2022). Tsunemoto (2023), according to Towell et al. (1996), classified mean length of run into breakdown fluency measurement as it incorporates pauses and may represent a speaker's hesitation. While Skehan (2003) put it under speed fluency measurement. Wright (2020) explained it as markers bounded by a silent pause.

Lastly, the location of pauses also needs to be taken into account. Mid-clause pauses may be employed for a variety of purposes, including the organisation of thought, the selection of appropriate vocabulary, the construction of sentence structure, and the emphasising of a word or phrase. End-clause pauses typically indicate the conclusion of a semantic unit and may also function as a cue to transition to a new topic or conclude the current one. Hawkins (1971) argued that clause-initial pauses were signs of content-related planning, while clause-internal pauses were likely to be symptoms of the speaker's search for lexis. While mid-clause pauses may impact the coherence of the verbal flow, end-clause pauses are a typical and expected interruption in the flow of speech. End-clause pauses are typically longer than pauses occurring in the middle of a sentence. This allows listeners sufficient time to process and comprehend the complete semantic unit.

Several of the above dimensions on pause are combined and calculated to produce various measures of breakdown fluency. Mid-clause pause duration measures the length of pauses that occur within a clause, which may reflect the speaker's difficulty in formulating language on the fly (Suzuki & Kormos, 2022). Huensch (2023) found that the only mid-clause pause measure that consistently distinguished between L1 and L2 speaking was the frequency of mid-clause pauses. Therefore, the study called for a more nuanced explanation of the link between mid-clause pauses and the stage of oral production, referring more to the related measures like mid-clause pause duration and mid-clause pause ratio (the average number of pauses between each sentence divided by the total number of syllables produced). Suzuki and Kormos (2022) also added the end-clause pause ratio into their measurement, which is calculated by dividing the number of pauses between clauses by the total number of syllables produced, providing an average measure of how frequently pauses interrupt the flow of speech at clause boundaries.

By examining these measures collectively, researchers can gain a more comprehensive understanding of an individual's speech fluency. It is important to note, however, that the perception of fluency is subjective and can vary among listeners (Bosker et al., 2014). Furthermore, cultural and linguistic backgrounds may also impact the utilisation and interpretation of pause phenomena (Jarvis, 2011).

In conclusion, the investigation of pause-related dimensions in speech fluency is multifaceted and requires a careful analysis of multiple factors. The combination of these variables facilitates a more precise evaluation of breakdown fluency, which is crucial for understanding the complexities of L2 acquisition and speech production.

2.2.3. Repair Fluency

As for repair fluency, compared to the first two fluency types, it has received less attention, and many scholars have chosen to perform statistics and calculations in a more general way. Tavakoli et al. (2020) measured the frequency of all repairs. Hadži et al. (2012) and Tsunemoto (2023) used a more general concept, the repair ratio, which is the total number of fluency dysfluencies (e.g., self-corrections and repetitions) divided by the total number of syllables. Wright (2013) proposed a concept that counts repetitions, retracings, repairs, and filled pauses and divides the number of them by the total number of characters. He named it the hesitation rate.

However, there are also more nuanced studies that subdivide repair fluency into self-correction, false start, and self-repetition ratio and then calculate their mean number separately (Suzuki and Kormos, 2022).

3. Dialogue Fluency Measurement

3.1. Dialogue Task Design

As mentioned above, monologue tasks have dominated research on L2 speech fluency, but there are still some studies that have used dialogue as a way of collecting speech samples. The speech samples in the study of Mora and Valls-Ferrer (2012) were all derived from the dialogue task. Participants were randomly assigned to work in pairs to complete the dialogue over the course. The dialogue format was a guided interview, with the presence of a researcher. The study employed a mutual questioning approach, wherein each participant was presented with a distinct set of seven printed questions pertaining to their university experience. Initially, one student posed all the questions to the other, after which the interviewer and interviewee alternated roles. The questions were considered to be of similar complexity, capable of eliciting similar language, and without significant differences in the dimensions of oral expression investigated. This design enhances the naturalness of the speaking data with respect to monologues and avoids the effects of dynamic interaction on fluency measures. However, it differs considerably from a real dialogue in that the two speakers are only questioners to each other, and there is no interaction

taking place. The dialogue task proposed by Skehan and Foster (2005) is more naturalistic in its approach. The task was in the form of a discussion in which participants were randomly divided into groups of two. They were provided with a list of people who had committed different crimes, and the participants discussed and then decided on the appropriate sentence for each case. The tasks were scheduled to take place during the regular class period, thus representing a realistic example of authentic classroom language learning. As the students were required to engage in discussion, the development of dialogue was unpredictable and the language used was spontaneous, bringing these dialogues closer to everyday communication.

In addition to studies that have employed dialogue alone as a research task, a growing number of studies have combined monologue and dialogue in order to provide a more comprehensive perspective on the study of L2 fluency development. In examining the effects of task repetition, Bygate (2001), in addition to using a film-cartoon retelling task, added a structured interview in which participants were asked to respond to questions about pictures reflecting different aspects of British life. This interview task involved spontaneous speech and relatively spontaneous interaction between speakers. However, such a guided approach to questioning may have an impact on the participants' expression. In contrast, Derwing et al. (2004) designed a task that allows for smoother transitions between monologue and dialogue tasks. The participants were required to complete three tasks, including both dialogues and monologues. The first task was a picture story, the second was an induced monologue on the theme of "the happiest moments of their lives", and the third was a dialogue based on the same theme. The latter could be described as a structured dialogue, but was more naturalistic than most samples of this type due to the interaction between speakers, the flexible timing of questions and answers, and the fact that the content came exclusively from the participants. This is a very clever arrangement, as the monologue topic can be used as an introduction, which saves the preparation time needed for the dialogue and allows the topic to flow very naturally, which will also make the participants more relaxed, thus making it easier for them to demonstrate their L2 speaking level. Wright's (2020) division of tasks is more nuanced. He compared the performance of participants in four tasks with different task loads (rehearsed vs. spontaneous speech, in monologic and dialogic mode). Participants completed the same two monologue speaking tasks and two dialogue tasks. All four tasks were based on everyday life topics that aligned with what was being taught in the classroom and were set up in such a way as to tap into testing participants' mastery of relevant and familiar vocabulary and grammatical structures. The tasks took the form of a final test at the participants' university, with the teacher playing the role of interlocutor in the dialogue tasks.

3.2. Dialogue Fluency Measurement

The aforementioned concepts pertaining to monologue measurement are equally applicable to dialogue and may be employed to measure within-turn fluency. A within-turn pause is where the first speaker produces the beginning of the turn and the second speaker completes it after a brief 0.4-second pause (Lerner, 1996). In view of the distinctive interactive nature of dialogue, a greater number of measures should be employed, as illustrated in Table 1 below (Peltonen, 2021; Wright, 2013; Witton-Davies, 2014).

Table 1. Dialogue fluency measures

Measurement	Definitions	Standardisation
Overlap duration	The speaking time of each participant during overlapping speech segments, which can indicate the level of interaction and potential competition for conversational space	Dividing the Overlap duration by the total dialogue time yields a percentage value indicating the proportion of overlap in the overall dialogue
Backchannel ratio	The mean number of lexical and non-lexical cues of words and noises from the listener, (divided by the total number of the total number of syllables produced)	It is already a ratio that can be used to directly compare the level of interaction between different dialogues or speakers
Turn pause duration	The time of pauses between each turn	To ascertain the average duration of each pause, the sum of all turn pause durations should be divided by the number of pauses. Alternatively, the sum of turn pause durations can be divided by the total dialogue time to determine the percentage of pauses in relation to dialogue.
Other-repetition number	The number of repetitions of words or longer utterances of the interlocutor's speech	Other-repetition number was calculated by dividing the total number of syllables in the dialogue by the number of repetitions per syllable, thus obtaining the average number of repetitions per syllable.
Collaborative completion number	The instances where a speaker completes an utterance started by a previous speaker, indicating a high level of cooperative interaction	To get the average probability of collaborative completion for each turn, divide the collaborative completion number by the number of turns in the whole dialogue.

4. Conclusion

This paper presents a relatively comprehensive review of research related to measuring L2 speech fluency from two different speech types: monologues and dialogues. The preceding analysis demonstrates that monologue is more common than dialogue in research related to L2 speech fluency. However, this has resulted in t discrepancies in the definition and operationalisation of the same measurement concept of monologue. Concerning dialogue, the paucity of

related studies precludes a systematic summary of the measurement methods employed by different scholars.

At the theoretical level, this study identifies deficiencies in the measurement of dialogues, to enhance the attention of the academic community on this aspect and advance the conceptualisation, operationalization, and measurement of dialogue fluency. In particular, the study seeks to provide definitions for the measurement of the unique interactional turns that appeared in dialogues and to standardise the process of measuring this type of measurement. With regard to the measurement of monologues, in light of the aforementioned standardisation discrepancies, further steps can be employed to ensure that data from disparate studies can be compared, thereby facilitating discussion and knowledge advancement.

From a pedagogical standpoint, an understanding of the utterance fluency measures influencing fluency perception enables educators to devise more efficacious pedagogical strategies, such as simulated dialogues and monologue exercises, to enhance students' speaking abilities. Furthermore, this study can guide curriculum design, assisting teachers in determining which speaking skills to prioritise and how to achieve an appropriate balance between fluency and other linguistic performance dimensions, such as accuracy and complexity. In terms of learning, students can learn to assess and enhance their verbal fluency, thereby enabling them to monitor and adapt their learning process and set realistic fluency goals. As for assessment, this systematic review can assist in the development and improvement of criteria for the assessment of L2 speech fluency, thereby ensuring greater fairness and consistency in the assessment process.

The research presented in this paper is limited by the small number of studies reviewed, which may not encompass all relevant data. Consequently, some findings that are representative of the field may have been overlooked. Furthermore, the studies may have failed to adequately take into account the differences in fluency across different cultural and linguistic contexts. Therefore, in terms of future research directions, it would be beneficial to enhance the breadth of research data collection in order to achieve a more comprehensive and systematic approach in terms of both quantity and time. It would be advantageous to conduct more fluency studies of various linguaculural contexts in order to gain a deeper understanding of how cultural differences affect fluency. Additionally, it would be valuable to develop and test new technological tools with the aim of more accurately measuring and analysing fluency.

References

[1] 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.

[2] Bosker, H. R., Quené, H., Sanders, T., & de Jong, N. H. (2014). The perception of fluency in native and non-native speech. *Language Learning, 64*(3), 579–614.

[3] De Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. Bilingualism: Language and cognition, 10(1), 7-21.

[4] 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.

[5] 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.

[6] Duez, D. (1985). Perception of silent pauses in continuous speech. Language and speech, 28(4), 377-389.

[7] Foster, P., & Ohta, A. S. (2005). Negotiation for meaning and peer assistance in second language classrooms. Applied linguistics, 26(3), 402-430.

[8] Foster, P. (2020). Oral fluency in a second language: A research agenda for the next ten years. Language Teaching, 53(4), 446-461.

[9] Freed, B. F. (1995). What makes us think that students who study abroad become fluent. Second language acquisition in a study abroad context/John Benjamins.

[10] Ginther, A., Dimova, S., & Yang, R. (2010). Conceptual and empirical relationships between temporal measures of fluency and oral English proficiency with implications for automated scoring. Language Testing, 27(3), 379-399.

[11] Hadži, V. P., Horga, D., & Bulc, T. B. (2012). Speech fluency: A result of oral language proficiency?. Linguistica, 52(1), 87-100.

[12] Hawkins, P. R. (1971). The syntactic location of hesitation pauses. Language and Speech, 14(3), 277-288.

[13] Huensch, A. (2023). Effects of speaking task and proficiency on the midclause pausing characteristics of L1 and L2 speech from the same speakers. Studies in Second Language Acquisition, 45(4), 1031-1055.

[14] Jarvis, S. (2011). Conceptual transfer: Crosslinguistic effects in categorization and construal. *Bilingualism: Language and Cognition, 14*(1), 1–8.

[15] Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. System, 32(2), 145-164.

[16] Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. Language learning, 40(3), 387-417.

[17] Lerner, G. H. (1996). On the" semi-permeable" character of grammatical units in conversation: Conditional entry into the turn space of another speaker. Studies in interactional sociolinguistics, 13, 238-276.

[18] Levelt, W. J. (1983). Monitoring and self-repair in speech. Cognition, 14(1), 41-104.

[19] Möhle, D. (1984). A comparison of the second language speech production of different native speakers. Second language productions, 26, 49.

[20] Mora, J. C., & Valls-Ferrer, M. (2012). Oral fluency, accuracy, and complexity in formal instruction and study abroad learning contexts. Tesol Quarterly, 46(4), 610-641.

[21] Peltonen, P. (2018). Exploring connections between first and second language fluency: A mixed methods approach. The Modern Language Journal, 102(4), 676-692.

[22] Peltonen, P. (2022). Connections between measured and assessed fluency in L2 peer interaction: a problem-solving perspective. International Review of Applied Linguistics in Language Teaching, 60(4), 983-1011.

[23] Pérez-Vidal, C., & Juan-Garau, M. (2011). The effect of context and input conditions on oral and written development: A study abroad perspective. International Review of Applied Linguistics in Language Teaching, 49, 157–185.

[24] 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.

[25] Sanaei, O., Zafarghandi, A. M., & Sabet, M. K. (2015). The effect of classroom anxiety on efl learner's oral narratives fluency: The case of intermediate level students. Theory and Practice in Language Studies, 5(7), 1390.

[26] Segalowitz, N. (2010). Cognitive bases of second language fluency. Routledge.

[27] Skehan, P. (2003). Task-based instruction. Language teaching, 36(1), 1-14.

[28] Skehan, P., & Foster, P. (2005). 7. Strategic and on-line planning: The influence of surprise information and task time on second language performance. In Planning and task performance in a second language, 193-216.

[29] Suzuki, S., & Kormos, J. (2022). The multidimensionality of second language oral fluency: Interfacing cognitive fluency and utterance fluency. Studies in Second Language Acquisition, 45(1), 38-64.

[30] Tavakoli, P., & Skehan, P. (2005). 9. Strategic planning, task structure and performance testing. Language learning & language teaching, 239-273.

[31] 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.

[32] Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. Applied linguistics, 17(1), 84-119.

[33] Tsunemoto, A., Uludag, P., McDonough, K., & Isaacs, T. (2023). The relationship between Japanese EFL learners' perceived fluency and temporal speech measures in a read-aloud task. JALT journal: the research journal of the Japan Association for Language Teaching, 45(1), 107-121.

[34] Witton-Davies, G. (2014). The study of fluency and its development in monologue and dialogue. Unpublished doctoral dissertation, Lancaster University, Lancaster, UK.

[35] Wright, C. (2013). An investigation of working memory effects on oral grammatical accuracy and fluency in producing questions in English. TESOL quarterly, 47(2), 352-374.

[36] Wright, C. (2020). Effects of task type on L2 Mandarin fluency development. Journal of Second Language Studies, 3(2), 157-179.