International Journal of English Language & Translation Studies
ISSN: 2308-5460

The Relationship between Temporal Measures of Oral Fluency and Ratings of Fluency: A Case of Iranian Advanced EFL Learners

[PP: 37-47]

Ali Akbar Farahani
Mohammad Hossein Kouhpaeenejad
(Corresponding Author)
Department of English Language and Literature
Faculty of Foreign Language and Literatures
University of Tehran, Tehran, **Iran**

ABSTRACT

Objective fluency judgment has always been a formidable task in language testing. Nonetheless, temporal fluency is the type of fluency which can be measured and quantified. Given that, temporal fluency is also known as temporal measures of fluency (Luoma, 2004). Furthermore, it has aroused considerable interest in analyzing speech of language learners in terms of quantitative measures (Kormos & Denes, 2004; Freed, 1995; Riggenbach, 1991; Lennon, 1990). They suggested that certain measures of fluency can more objectively specify fluency level and that perceptual understanding of fluency to a high extent correlate with these measures. Following these studies, the present study was an endeavor to relate quantitative measures of fluency and assessment of fluency in oral speech of L2 learners. To do so 30 advanced EFL learners whose speaking score on TOEFL iBT scale was between 19 to 22, i.e. B2 on CEFR scale, were selected. Then, they were given a picture strip as the elicitation task and asked to make up a story based on that. Their voice was recorded, transcribed and further analyzed by voice analysis software called PRAAT to calculate seven measures of fluency. Meanwhile, two trained listeners were required to rate the recordings, scoring them from 1 to 9. Finally, the relationship between these variables was calculated. The results showed that judge listeners' ratings of fluency were highly correlated with speech rate, phonation time ratio, and mean length of runs. Moreover, among the measures of temporal fluency speech rate proved significantly correlated with articulation rate, phonation time ratio, and mean length of runs.

Keywords: *Temporal Fluency, Language Testing, PRAAT, Speech Rate, Iranian EFL Learners*

ARTICLE INFO	The paper received on	Reviewed on	Accepted after revisions on
	21/05/2017	**12/06/2017**	**19/08/2017**

Suggested citation:
Farahani, A. & Kouhpaeenejad, M. (2017). The Relationship between Temporal Measures of Oral Fluency and Ratings of Fluency: A Case of Iranian Advanced EFL Learners. *International Journal of English Language & Translation Studies.* 5*(3). 37-47.*

1. Introduction

Fluency may be one of the most common terms used in a wide variety of senses in English language teaching and testing. To clarify the point, it is sometimes claimed that one can speak English fluently or the other is a fluent speaker of English but it is not clear to what extent they master the language. However, Fillmore (1979) defines fluency as the speaker's ability to fill time with talk, and when speakers are fluent in this way, they do not have to stop many times to decide on what to say next or how to formulate it. He further explains that fluency depends on a variety of factors such as quick access to a wide range of words and practiced control over syntactic devices. Simply put, fluency is the ability to promptly decide when it is appropriate and efficient to use lexicon.

In a similar vein, Leonard and Shea (2017) define fluency as "the temporal characteristics of speech, including such aspects as pausing, speed (speech rate), and repair (how often speakers make false starts or self-corrections)" (p. 2).

Even with Fillmore's definition at hand, it still seems virtually impossible to avoid misjudgment of one's L2 speech performance due to the lack of standardized assessment tools, leading to subjective and hence unreliable decisions at times. Numerous studies have been conducted in order to introduce a more refined definition of fluency and to formulate appropriate assessment criteria, which can in turn add to objectivity of fluency judgment. Among those is a comprehensive study on Hungarian English L2 learners by Kormos and Denes (2004) which also initially motivated the design of this study which is

focused on a group of 30 Iranian advanced learners of English as a foreign language whose fluency as a temporal phenomenon in their L2 oral performances was rated by judge listeners.

This study is different from other studies in that they were all carried out in ESL context, while this one was carried out in EFL context. It goes without saying that contrary to Europeans who can easily access native speakers and other foreign language resources as a result of a more cosmopolitan atmosphere and easier global mobility, Iranian learners of foreign languages' exposure to language input is limited to a few hours of classroom teaching and teachers' oral output. Additionally, against most languages spoken in Europe the alphabet and left to right writing system of which resemble those of English, Farsi has completely different alphabet and witting system.

2. Review of the Literature

2.1 Fluency

An overarching account of fluency, which is one of the most controversial terms in both applied linguistics and SLA, has always eluded the researchers. This seems to be the reason why it has been discussed in the literature from a wide variety of perspectives. Yet, researchers in this area have tried to come up with their own definitions: "the ability to produce continuous speech without causing comprehension difficulties or a breakdown of communication" (Richards & Schmidt, 2002) or "When a language is fluent, it is spoken easily and without many pauses" (Cambridge advanced learner's dictionary) or as Harrell (2007) puts it "a speech language pathology term that means the smoothness or flow with which sounds, syllables, words and phrases are joined together when speaking quickly"(p. 65).

According to Harrell (2007), fluency is used in an informal way to represent a high level of language expertise in a foreign language or another learned language. Koponen (1995), however, defines fluency with reference to flow or smoothness of speech, rate of speech, absence of excessive pausing, absence of disturbing hesitation markers, length of utterance, and connectedness of speech. Fillmore (1979) classifies fluency in terms of scope so that in the first category which is a "broad" one, fluency includes a number of components such as pausing, complexity, coherence, appropriateness, and creativity. On the other hand, in the second category that is a "narrow" one, fluency is defined as normal flow of speech. In communicative language teaching, fluency is defined as "natural language use occurring when a speaker engages in meaningful interaction and maintains comprehensible and ongoing communication despite limitations in his or her communicative competence" (Richards, 2006, p.14). The word 'fluency', has a Latin origin meaning 'flow'. 'Fluency' in many other languages has similar meanings such as flow and fluidity (Koponen & Riggenbach, 2000). The definitions of the term in applied linguistics also seem to have at least one feature like 'fluidity' in common.

Fillmore (1979, as cited in Kormos, 2006) points four different interpretations out: 1) The ability to talk at length with few pauses and fill time with talk 2) The ability to express message in a coherent, reasoned and "semantically dense" manner 3) The ability to know what to say in a wide range of contexts 4) The ability to be creative and imaginative in language use. As a highly fluent speaker, according to Fillmore has all the abilities mentioned above. This definition is one of a few early definitions which include both qualitative and quantitative aspects. Moreover, although L2 learners are not considered in Fillmore's definition, Fillmore (1979) distinguishes between fluency in monologues and dialogues in that a wide vocabulary used in monologues would enhance speaker's fluency while vocabulary size does not play this decisive role in dialogues in which other abilities of speakers (e.g. the ability to follow the conversation) count (Mizera, 2005). Thus the speakers' fluency in monologues would be higher than their fluency in dialogues. Notwithstanding this effective role, the fourth interpretation of fluency by Fillmore is more valued in dialogues in which speakers have limited control over the topic. The interpretation of fluency is the ability to "fill time with talk" demonstrates the significance he attached to it, though not clearly on formulaic expressions' role in achieving oral fluency. This essential role of formulaic expressions has been also stressed in a number of studies investigating fluency in L2 speech (e.g. Ejzenberg, 2000; Towell et al.,1996). Ejzenberg's (2000) study is a case in point. Ejzenberg investigated the use of formulaic language among fluent and non-fluent speakers. The results highlighted the ability of fluent speakers in using chunks

appropriately compared to non-fluent speakers who fail to do this.

Formulaic Language has been defined as "sequences of words that are stored and retrieved as a unit from memory at the time of use, rather than generated online using the full resources of the grammar of the language." (Richards & Schmidt, 2002, p.210). Based on this definition, retrieving cluster of words places less demand on memory than producing novel linguistic structures. Thus, the speaker can produce words more quickly, hence speaking more fluently. Although the role of formulaic language in enhancing fluency is acknowledged (Wood, 2006), Rehbein (1987) believes that "speech formulae can also prevent learners from developing native-like fluency" (p.104). He mentions a native speaker's judgment to support his claim.

2. 2 Measures of Oral Fluency

When it comes to empirical studies on fluency, as Kormos and Denes (2004) discuss, researchers have adopted three different approaches:

First, the development of fluency has longitudinally been investigated (Freed, 2000; Huensch, & Thompson, 2017, Lennon, 1990; Towell et al.,1996)

Second, fluent and non-fluent speakers have been compared (Ejzenberg, 2000; Tonkyn, 2001). Third, fluency scores with temporal variables are correlated (Fulcher, 1996).

However, the common thread running through all of them is that the best predictors of fluency are *speech rate*, that is, the number of syllables articulated per minute and the mean length of runs, that is, the average number of syllables produced in utterances between pauses of 0.25 seconds and above (e.g. Ejzenberg, 2000; Freed, 1995, 2000; Lennon, 1990; Riggenbach, 1991, Towell et al, 1996); *Phonation-time ratio*, that is, the percentage of time spent speaking as a percentage proportion of the time taken to produce the speech sample, has also been pointed out to be a predictor of fluency (Towell et al, 1996; Lennon, 1990).

Most researchers agree that disfluencies tend to occur in clusters in the speech of non-fluent L2 learners (e.g. Freed, 1995, 2000; Riggenbach 1991), while fluent students tend to pause at grammatical junctures (Lennon, 1990; Towell et al., 1996). Fulcher (1996) concluded that low-proficiency students tend to hesitate because they have problems retrieving lexical items, encoding the grammatical form of their message and correcting their own output. On the other hand, high-proficiency students are able to plan in advance and mostly hesitate only when they want to express complex ideas.

The common European Framework of Reference (CEFR), in the same line, has introduced a set of descriptors for spoken fluency:

Table: 1 Descriptors for spoken fluency, CEFR Manual

	SPOKEN FLUENCY
C2	Can express him/herself at length with a natural, effortless, unhesitating flow. Pauses only to reflect on
C1	Can express him/herself fluently and spontaneously, almost effortlessly. Only a conceptually difficult
B2	Can communicate spontaneously, often showing remarkable fluency and ease of expression in even longer complex stretches of speech
	Can produce stretches of language with a fairly even tempo; although he/she can be hesitant as he/she searches for patterns and expressions, there are few noticeably long pauses. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without imposing strain on either party.
B1	Can express him/herself with relative ease. Despite some problems with formulation resulting in pauses and 'cul-de-sacs', he/she is able to keep going effectively without help.
	Can keep going comprehensibly, even though pausing for grammatical and lexical planning and repair is very evident, especially in longer stretches of free production.
A2	Can make him/herself understood in short contributions, even though pauses, false starts and reformulation are very evident.
	Can construct phrases on familiar topics with sufficient ease to handle short exchanges, despite very noticeable hesitation and false starts.
A1	Can manage very short, isolated, mainly pre-packaged utterances, with much pausing to search for expressions, to articulate less familiar words, and to repair communication

What adds to the difficulty of objective evaluation of L2 learner's oral speech is mixing the quantitative aspects of fluency descriptors such as 'pauses' and 'false starts' with qualitative features like 'relative ease' and 'fairly even tempo'.

Having assumed that fluency is context-dependent (e.g. Rehbein, 1987; Sajavaara, 1987; Lennon, 1990), Riggenbach (1991) delved into the analysis of temporal variables underlying second language fluency with the investigation of interactive features. She concluded that topic initiations, backchannels, substantive comments, latching and overlapping as well as the amount of speech produced also contributed to fluency judgments, though to a limited extent.

As for phonological research, Hieke (1985) established additional measures of fluency on the basis of the presupposition that fluent speech equals connected speech, in which certain phonological procedures, such as consonant attraction are at work. Consonant attraction "occurs where final consonants are drawn to the following syllable if that begins with a vowel" (Hieke, 1985, p. 140). In an earlier study, Hieke (1984) found that consonant attraction can be a reliable indicator of the fluency of non-native speech in informal English style. Moreover, Wennerstorm (2000) in her research investigated how intonation influences the perception of fluency by means of analyzing dialogues between speakers of English as a second language

and native English speakers. Her study concludes that it is the ability to speak in phrases instead of speaking word by word that can lead to the perception of fluent speech, rather than longer utterances or shorter pauses.

Vanderplank (1993), in another study, suggests that pacing (the number of stressed words per minute) and spacing (the proportion of stressed words to the total number of words) are better indicators of difficulty in listening materials than standard speech rate measures such as syllable per minute. This would mean that these variables are also useful in predicting fluency scores. Towell et al (1996) investigated what qualitative changes take place in the use of formulaic language parallel to the increase of fluency after participants spent a year in the target language environment. They found that the two selected students improved in how they employed different types of formulae after their stay abroad. Ejzenberg (2000) compared how fluent and non-fluent speakers employ formulaic language. Her results also showed that fluent students were able to make use of prefabricated chunks more efficiently, whereas non-fluent learners frequently used formulae inappropriately.

This study is different from other studies in that they were all carried out in ESL context, while this one was carried out in EFL context. It goes without saying that contrary to Europeans who can easily access native speakers and other foreign language resources as a result of a more cosmopolitan atmosphere and easier global mobility, Iranian learners of foreign languages' exposure to language input is limited to a few hours of classroom teaching and teachers' oral output. Additionally, against most languages spoken in Europe the alphabet and left to right writing system of which resemble those of English, Farsi has completely different alphabet and witting system.

2.3 Temporal Measures of Fluency

As Freed (1995) points out, the concept of fluency hinges upon temporal aspects of speech such as speaking rate, speech-pause relationships, and fluency of dysfluency markers like hesitation, repetition and self-correction measured by machine or by human impression. The Chambers' (1997) position can be a good point of departure in this regard, hence providing sufficient grounds for the temporal measures of oral fluency:

A definition restricting fluency in spoken production to temporal variables, such as pauses of various kinds and length of runs between pauses provides a useful anchorage for a concept which is prone to vagueness and multiple interpretations. Temporal variables in speech production are empirically identifiable and quantifiable. The study of temporal variables also enables psycholinguistic research to gather valuable empirical evidence since processes of language production themselves are not directly accessible. Whereas appreciating a skill is a qualitative judgment (one is reminded of the mark for artistic interpretation in ice-skating implied by terms such as "smoothness" or "ease"), a performance in real time has quantifiable aspects such as rate of speech, frequency and location of silences and hesitations (Chambers, 1997; p.538).

Temporal fluency is the type of fluency which can be measured and quantified. Given that, temporal fluency is also known as temporal measures of fluency (Luoma, 2004). Like perceptual fluency which is useful in assessing oral fluency, temporal fluency, as a set of measurable variables, can also be considered useful for this purpose. As a general rule of thumb, the researchers in this area would agree that no other variable in an individual's spoken output is as empirically identifiable and quantifiable as temporal variables. These are possibly the most distinctive variables that psycholinguists have at their disposal to investigate speech production. As a result, the studies on fluency as a temporal phenomenon would result in more practical approaches to study of speech production and similar areas within psycholinguistics and second language development. It is worth noting that temporal fluency is often quantified on the basis of the number of words or syllables spoken or the number or the lengths of hesitation pauses inserted in the delivery (Wood, 2012).

Kang (2008) classifies temporal measures of fluency in two main categories:
1. Rate measures, including a) Speech rate b) Articulation rate c) Phonation time ratio d) Mean length of runs
2. Pause measures, including a) Mean length of pauses b) Number of silent pauses per minute c) Number of filled pauses per minute

Kormos (2006) lists most frequently studied measures of fluency along with their definitions.

Table: 2 Measures of Fluency by Kormos (2006)

Measures	Definitions
Speech rate	The total number of syllables produced in a given speech sample divided by the amount of total time required to produce the sample (including pause time) expressed in seconds. This figure is then multiplied by sixty to give a figure expressed in syllables per minute. Riggenbach (1991) suggested that unfilled pauses under 3 seconds should not be included in the calculation of speech rate.
Articulation rate	The total number of syllables produced in a given speech sample divided by the amount of time taken to produce the min seconds, which is then multiplied by sixty. Unlike in the calculation of speech rate, pause time is excluded. Articulation rate is expressed as the mean number of syllables produced per minute over the total amount of time spent speaking when producing the speech sample.
Phonation-time ratio	The percentage of time spent speaking as a percentage proportion of the time taken to produce the speech sample (Towell, Hawkins, & Bazergui, 1996).
Mean length of runs	An average number of syllables produced in utterances between pauses of 0.25 seconds and above.
The number of silent pauses per minute	The total number of pauses over 0.2 sec divided by the total amount of time spent speaking expressed in seconds and is multiplied by 60.
The mean length of pauses	The total length of pauses above 02 seconds divided by the total number of pauses above 0.2 seconds.
The number of filled pauses per minute	The total number of filled pauses such as uhm, er, mm divided by the total amount of time expressed in seconds and multiplied by 60.
The number of disfluencies per minutes	The total number of disfluencies such as repetitions, restarts of and repairs are divided by the total amount of time expressed in seconds and multiplied by 60.
Pace	The number of stressed words per minute (Vanderplank,1993)
Space	The proportion of stressed words to the total number of words (Vanderplank, 1993)

In another categorization, Skehan (2003) has distinguished between four types of fluency:

1. Breakdown fluency (silence)
2. Repair fluency ("reformulation, replacement, false starts, and repetition")
3. Speech rate (speed fluency)
4. Automatization ("through measures of length of run") (p.8). Notwithstanding differences, such categorizations have some measures of frequency in common.

In a recent study, Huensch and Tracy-Ventura (2017) investigated the effect a period of residence abroad on different aspects of fluency. Their results in indicated that speed fluency was more quickly improved and was less prone to attrition after returning home. On the other hand, breakdown fluency was less affected by residence in the L2 context and was more prone to attrition after returning home. Interestingly, there were no effects on repair fluency at all. Hernandez (2016) also reports similar results.

Oral fluency along with its relationship to temporal measures has received some attention in the related literature. However, it is still not clear how different measures of L2 fluency correlate with the judges' ratings of fluency with respect to the common threads which possibly run through them. Taking the legacy left by the pioneering works in the realm of fluency judgment, the present study is an attempt to shed some light on such areas in order to offer insights into the evaluation procedures for judging oral fluency of EFL learners. Specifically, the following research questions are posed:

Is there any relationship between temporal measures of fluency and the judges' ratings of fluency in L2 oral speech?

Which temporal measures of fluency do significantly correlate with one another?

3. The Present Study

3.1 Participants

A total of 30 male (n = 15) and female (n = 15) Iranian learners of English as a foreign language, aged from 22 to 30, participated in this study. They were all university graduates attending the University of Tehran Language Center, Building no. 3 to prepare for the TOEFL test. The participants were then required to take the placement test which contained printed questions of TOEFL iBT for reading, listening, and writing (See Appendix B). The speaking test was conducted as a 7-to-9 minute interview. Among them, those who scored between 75 to 90 out of 120, with their speaking scores ranging from 19 to 22, i.e. B2 on CEFR scale, were chosen for the recording task. Moreover, like any other Iranian student holding a bachelor's degree, they had also done three English courses during 4 years on the two-hour-a-week basis. Nonetheless, it is worth noting that despite the time spent on English language education at university as well as school, the teaching approach and course books are not effective enough to prepare students for the future communication specifically in terms of oral proficiency.

Iranian students, though keen on speaking English outside classroom, have limited opportunities, for the country's atmosphere is not as international as it should be for a variety of reasons, meaning that students' exposure to English would be mainly through American movies

The main reason behind selection criteria was to control as many participant variables as possible including: education, socioeconomic setting, language learning background and current language environment, and level of L2 spoken proficiency.

In this study, 2 non-native speakers of English participated as judges, who were both males. They were both graduates in TEFL from university of Tehran in Iran. They were teaching at the language center of the University and had several years of experience in assessing oral proficiency of the English L2 learners. All the cooperation on both participants and judges' side was voluntary.

3.2 Instruments

3.2.1 TOEFL iBT test

A real TOEFL iBT test was administered to check students' proficiency level at their entrance to the institutes TOEFL classes. Unlike the real TOEFL iBT, the test is given to the applicant in paper, including 4 skills. The test starts with an hour of reading comprehension including 3 reading passages, each with 14 questions, followed by 55 minutes listening comprehension with 6 listening passages and an overall of 34 questions. After that comes the writing section containing one essay question requiring students to write an essay of at least 300 words long in 30 minutes. Finally, all the applicants are interviewed by trained TOEFL instructors for about 7 to 9 minutes.

3.2.2 Picture strip

A cartoon, as a picture description device, consisting of 6 pictures in logical order (See appendix A) was used to elicit the speech data. It was extracted from "Vater und Sohn", a book by the German artist, Erich Ohser. The choice of the cartoon over a reading task was based on the interpretability of the story and easiness of the vocabulary needed to describe it. As Riazantseva (2001, p.506) notes, "the cartoon description is a highly structured task, as it offers minimal freedom of choice (grammatical, lexical, and semantic)". Additionally, compared with a reading task, a picture description task reduces hesitations caused by reading effects (coding) in readers' speech.

In this study, unlike plenty of the previous ones in which the participants were given two or three sets of cartoons to choose from, the participants were given only one set of cartoon and were then asked to make up a story for it. This would naturally facilitate the arduous task of flouncy assessment by judges, leading to higher reliability.

3.2.3 PRAAT

PRAAT is computer software used for analyzing speech and distinguishing silent pauses from phonations through providing oscillographic pictures in which silent pauses are separated from phonations. Such pictures were generated by the software to measure the lengths of pauses (see Figure 1). In the oscillographic pictures, silent pauses are mainly shown by straight and flat portions of the line, while sounds, whether they are vocal or background, are represented by wavy portions of the line.

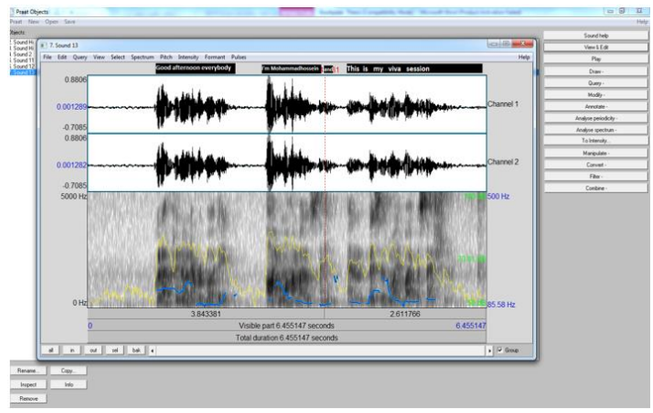
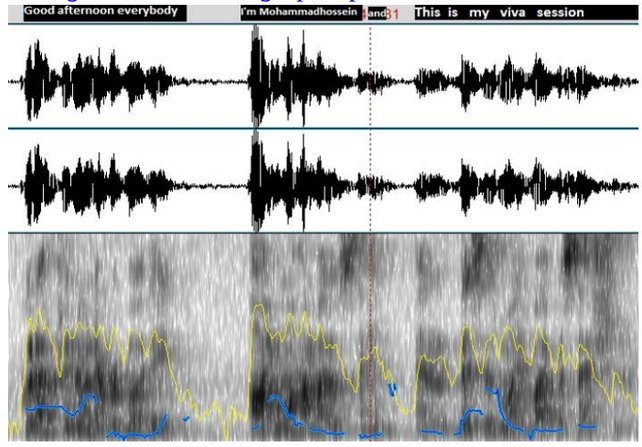


Figure 2 provides a clearer oscillographic picture in which this sentence was spoken: "*Good afternoon everybody. I'm Mohammadhossein, and this is my viva session."* In this picture silent pauses are shown in light gray color in the lower part of the picture or are flat as seen in the upper part.

Figure: 2 An oscillographic picture on PRAAT



However, there is one problem in which foils the attempt to distinguish silent pauses from utterances just by looking at the oscillographic pictures, which is the possibility of mistaking silent pauses for utterances, because it is not clear from the graph whether the vertical lines indicate vocal sounds or silent pauses, for silent pauses include sounds for breathing which are often shown by vertical lines just like vocal sounds on the graph. Therefore, to avoid such confusion the researcher is required to both visually identify vertical lines on the graph and listen to the recorded sounds to differentiate silent pauses from utterances.

As for listening to the recordings, the researcher can highlight one part of the line by simply dragging the cursor on the part in the graph, and then click on the highlighted bar to play the sound of the part, which would enable him/her to concentrate on that part to distinguish the nature of the sound. PRAAT also allows the user to magnify the sound on the graph and replay

Figure: 1 A snapshot of the software PRAAT

Cite this article as: Farahani, A. & Kouhpaeenejad, M. (2017). The Relationship between Temporal Measures of Oral Fluency and Ratings of Fluency: A Case of Iranian Advanced EFL Learners. *International Journal of English Language & Translation Studies.* 5*(3). 37-47.*

it in case he/she has difficulty identifying a sound.

3.3 Procedure

Participants who agreed to do the recording task at the presence of the researcher were given a hard copy of the cartoon strip, and those who preferred to do the recording at home were sent a soft copy of the cartoon strip via email. They were then required to spend 2 minutes looking at the picture and start telling the story while recording their voice session separately at their home in a very quiet room.

Digital audio recorders, cellphone, and laptop were used by the researcher and participants for the recording task. The participants recoded their voices with no interruption or help from the researcher or a third party.

After collecting the data, the participants' recordings were carefully listened to and transcribed. The transcription was done both by the researcher and a number of participants. The number of syllables in each speech sample was counted manually using the transcripts. Then, using PRAAT, the researcher measured each silent pause in millisecond, and analyzed the data for temporal variables using PRAAT software. According to Riazantseva (2001, P.508 citing Duez, 1982), silent pause was defined as "any interval of the oscillographic trace where the amplitude is indistinguishable from that of the background noise". After that, through the following mathematical formulas and based upon the total response time, 7 temporal measures of fluency, which were discussed in the first chapter (table 1.1), were calculated.

1. Speech Rate (SR):

$$SR = \frac{\text{The total number of syllables produced in a given speech sample}}{\text{Total time taken to produce the speech sample (in seconds) including pause time}} \times 60$$

In this study, "Speech rate", as the most important fluency variable, is used as pruned speech rate (Lennon, 1990) that is the rate of produced syllables excluding repetitions and corrections. Moreover, contrary to Riggenbach's (cited in Kormos et al, 2004) suggestion, all pauses including both under or over 3 seconds were considered when calculating of total time of speech sample. Speech rate is expressed in syllables per minute.

2. Articulation Rate (AR):

$$AR = \frac{\text{The total number of syllables produced in a given speech sample}}{\text{Total time taken to produce the speech sample (in seconds) excluding pause time}} \times 60$$

According to Kormos et al (2004, p.152 citing Riggenbach, 1991) "Pauses shorter than 0.2 seconds are considered micropauses and are not regarded as hesitation phenomena." Therefore, pauses under 0.2 were not excluded from the amount of total time. Articulation rate, like speech rate, is expressed in syllables per minute.

3. Phonation-Time Ratio (PTR):

$$PTR = \frac{\text{Time spent speaking (without pause time)}}{\text{Total time taken to produce the speech sample (in seconds)}} \times 100$$

Phonation time is expressed in percentage. Regarding mathematical relation between SR and AR, dividing speech rate by articulation rate also gives the phonation-time ratio:

$$PTR = \frac{\text{Speech rate}}{\text{Articulation rate}} \times 100$$

4. Mean Length of Runs (MLR)

$$MLR = \frac{\text{Total number of syllubles in all runs}}{\text{Total number of runs}}$$

Mean length of run is of paramount importance since it indicates that that "fluent speech involves the use of a large repertoire of formulaic sequences to aid in balancing skills, attention, and planning during spontaneous speech" (Wood, 2007, p. 211). A run is defined as an utterance produced between pauses of 0.25 seconds and above (Towell et al, 1996). MLR is expressed in number of syllables.

5. Mean Length of Pauses (MLP)

$$MLP = \frac{\text{Total length of pauses above 0.2 seconds}}{\text{Total number of pauses above 0.2 seconds}}$$

As discussed for calculation of articulation rate, pauses shorter than 0.2 are not regarded as hesitation so they're not included in total length of pauses.

6. The Number of Silent Pauses Per Minute (NSPPM)

$$NSPPM = \frac{\text{The total number of pauses above 0.2 second}}{\text{Total time taken to produce the speech sample (in seconds)}} \times 60$$

Following Riggenbach, the pauses shorter than 0.2 are considered as micro-pauses and are excluded from the calculation.

7. The Number of Filled Pauses per Minute (NFPPM)

$$NFPPM = \frac{\text{The total number of filled pauses}}{\text{Total time taken to produce the speech sample (in seconds)}} \times 60$$

Filled pauses are silences filled by gap fillers such as uhm, er and mm.

Following data collection and the above-mentioned calculation, the

recordings together with a speech evaluation form (See appendix A) were given to the judge listeners, and they were then asked to rate the oral performances on a nine-point scale (1= extremely dysfluent, 9= extremely fluent). All the judges had already been briefed on the purpose of study and scoring procedure. They were also asked to contact the researcher if needed.

4. Results and Discussion

As mentioned in chapter 3, PRAAT, the voice analysis software, was used to investigate these questions by objectively measuring the 7 temporal measures of fluency outlined in the previous chapter. In what follows the findings of this investigation are presented. The descriptive statistics for the seven measures of fluency are displayed in Table 3.

Table:3 Descriptive Statistics of Seven Measures of Fluency

	N	Minimum	Maximum	Mean	Std. Deviation
Speech Rate	30	45.22	126.27	89.6057	24.74413
Articulation Rate	30	53.45	143.55	92.7310	25.43520
Phonation-time Ratio	30	33.60	96.80	60.2877	18.11171
Mean length of runs	30	2.30	18.40	9.2780	4.58355
Mean length of pauses	30	.60	2.73	1.2869	.68068
Silent pauses per minute	30	5.40	23.90	12.6380	6.97157
Filled pauses per minute	30	4.19	13.80	8.1193	3.18719

Considering fluency a temporal phenomenon, it is hypothesized that temporal features of fluency are highly likely to correlate with trained listeners' perception of fluency. The relations between temporal variables are also of significance to the researcher as they may either reveal or even deny the temporal nature of oral fluency.

Correlations between temporal measures and scores of fluency

This part is mainly focused on the first research question in which the relationship between different trained listeners' scores and temporal measures of Iranian learners' performance was to be investigated. Table 4 displays the correlations between the judges' ratings and measures of fluency.

Table: 4 Correlations among the Judges' Ratings and Measures of Fluency

	Speech rate	Articulation rate	Phonation-time ratio	Mean length of runs	Mean length of pauses	NSP per minute	NFP per minute
Correlation	.607**	.609**	.415*	.623**	-.454*	-.394*	-.020
Sig.	.000	.000	.022	.000	.012	.031	.919

**. Correlation is significant at the 0.01 level (2-tailed)

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

Note. NSP = number of silent pauses. NFP = number of filled pauses

As seen in the table, the ratings are significantly correlated with speech rate, articulation rate, phonation time ratio, mean length of runs and mean length of pauses, and moderately with number of silent pauses per minute. Among them the highest correlations are with speech rate, articulation rate, and mean length of runs, with r being .60, .60, and .62 respectively. These correlations are positive and these measures are mainly based on utterances (i.e., number of words or syllables), an indication of fluency; however, the correlations with Mean Length of Pauses and Silent Pauses per Minute are negative. The ratings also have near zero correlations with the last measure of fluency (i.e., number of filled pauses per Minute).

Correlations between temporal variables of fluency

This part addresses the second research question regarding the relationship between different temporal measures fluency. In chapter 3 the formulas and the way of calculating these measures were outlined. Likewise, they are again discussed here, but in more details.

It appears from Table 5 that all correlations among the first five measures are significant although the correlations of the mean length of pauses with the other four measures are all negative. These measures also have negative or near zero correlations with the last two measures. In general, it appears that the last two factors do not have much common variance with the first five measures. The interpretation would be they are not measuring the same construct.

Table: 5 The Correlations Matrix for Measures of Fluency

		Speech Rate	Articulation Rate	Phonation-time Ratio	Mean length of runs	Mean length of pauses	Number of Silent pauses per minute	Number of Filled pauses per minute
Speech Rate	Pearson Correlation	1	.867**	.652**	.734**	-.643**	-.311	-.168
	Sig. (2-tailed)		.000	.000	.000	.000	.095	.375
Articulation Rate	Pearson Correlation			.573**	.611**	-.558**	-.221	.028
	Sig. (2-tailed)			.001	.000	.001	.240	.885
Phonation-time Ratio	Pearson Correlation				.477**	-.384*	-.104	.026
	Sig. (2-tailed)				.008	.036	.583	.892
Mean length of runs	Pearson Correlation					-.387*	-.376*	.084
	Sig. (2-tailed)					.035	.041	.658
Mean length of pauses	Pearson Correlation						.192	.079
	Sig. (2-tailed)						.309	.680
Silent pauses per minute	Pearson Correlation							-.104
	Sig. (2-tailed)							.584

**. Correlation is significant at the 0.01 level (2-tailed).

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

Speech Rate (SR) & Articulation Rate (AR)

Speech Rate (SR) is calculated through dividing total number of syllables uttered by the total time taken *including* pause time. Articulation rate (AR), on the other hand, is measured by dividing the number of syllables uttered by the total time taken *excluding* pause time. Total silent pauses time was subtracted from the total response time in order to calculate total time of phonation or articulation. Compared to articulation rate, the values of speech rate are smaller, for the denominator of articulation rate (i.e. phonation time) is smaller than that of speech rate (i.e. total response time including both phonation time and silent pause time). Therefore, a more fluent speaker in terms of speech rate has to use both more syllables for utterances and shorter time for pauses. Even if the speaker produces more syllables, it does not necessarily mean that the speaker's produced syllables on this measure is higher because the speaker's time for pau**ses might be longer. As table 5** displays, speech rate and articulation rate are highly correlated. (r=0.867; p= 0.01)

Speech Rate (SR) & Phonation Time Ratio (PTR)

Phonation time ratio, which is solely based on temporal factors, is the amount of time spends speaking as a percentage of time taken to produce the speech sample (Towel et, al, 1996). If a speaker uses pauses that reach 20 percent of the total response, then his / her Phonation time ratio is 80 percent. In order to be a fluent speaker in terms of phonation time ratio, how fast a speaker 'articulates' utterances does not matter; only ratio of phonation time and silent pause time matter. According to table 5 the correlation between speech rate and phonation-time ratio is significantly positive($r = 0.65$; $p < 0.01$).

Speech Rate (SR) & Mean Length of Runs (MLR)

Mean length of run is defined as the mean number of syllables produced between hesitations longer than, in this study, 0.25 seconds, meaning that when a fluent speech run includes a.24 second pause, the run is still considered one run in this study. The weakness of this measure is that different cut-offs for pauses would lead to different results. The results show a significant positive correlation of 0.73 between speech rate and mean length of run.

Phonation Time Ratio (PTR) & Mean Length of Runs (MLR)

The definitions and calculations method of phonation time ratio and mean length of runs were explained above. These two measures of fluency, as displayed in table **5**, proved significantly correlated (r=0.47).

Mean Length of Pauses (MLP) & Speech Rate (SR)

Mean length of pauses (MLP) is the average length of pauses that are longer than.25 seconds and is calculated through dividing total length of pauses above .2 seconds by total number of pauses above .2 seconds. As seen in table 5, there is a negative correlation between mean length of pauses and speech rate. (r= -.64). According to Ushigusa (2008) what makes MLP important in judging fluency is that even if MLP is constant between two speakers, one of the speakers might be more nonfluent than the other, for they can use pauses more frequently and sound less fluent than the other.

Considering the results, the research questions posed in this paper are answered individually.

1. Is there any relationship between temporal measures of fluency and the judges' ratings of fluency in L2 oral speech?

Average fluency score of participants, given by trained listeners was highly correlated with three temporal measures: speech rate, articulation rate, and mean length of runs. These high correlations of speech rate, articulation rate, and mean length of runs with fluency score make these measures the most salient predictors of fluency judgments. The findings are in line with the result of many other studies (e.g. Ejzenberg, 2000; Kormos et al, 2004; Lennon, 1990; Tower et al, 19960. The present study also found that the other two measure of fluency namely mean length of pauses and number of pauses which are specified employing length and number of pauses are not good indicators of fluency, but disfluecy.

2. Which temporal measures of fluency do significantly correlate with one another?

The study also found close relationships between four temporal measures of fluency, making them good predictors of fluency scores: Speech rate, articulation rate, phonation-time ratio and mean length of runs. Pausing measures such as number of filled/silent pauses per minute or did not show significant correlations with any of those four measures or judges' scores. However, mean length of pauses

was negatively correlated with speech rate and articulation rat.

5. Conclusion

The present research was carried out to explore the relationship among temporal measures of fluency as a component of oral proficiency in speech of 30 Iranian L2 learners of English. The study also took account of perception of fluency by trained listeners and its correlation with temporal measures. To do so, the design of this study was led by two research questions investigating correlations in two groups of variables:

- Between temporal measures and fluency scores
- Between temporal measures

According to Wood (2012), in most studies speech and articulation rate increase with overall fluency or correlate with evaluation of fluency. The findings of the current study confirm Wood's claim. In the same line, the results also lend support to the outcome of other studies such as the one by Lennon (2000) in which the speed aspect of fluency definitions was underlined, as discussed in literature review: "a working definition of fluency might be the rapid, smooth, accurate, lucid, and efficient translation of thought or communicative intention into language under the temporal constraints of on-line processing". (Lennon, 2000, p.26)

Articulation Rate which, based on Wood (2012), is "fairly a sound indicator" illuminates how fast learners produced utterances while they were saying those utterances. This section has examined how many syllables they produced per 60 seconds of utterances. However, this measure does not consider the time for pausing to think about what to say. No matter how fast a speaker 'articulates' utterances, the speaker might sound nonfluent if he / she uses many and / or long pauses between those utterances.

Towell et al. (1996) elucidated the participant's improvement of oral fluency identified in their increased values of the temporal measure known as mean lengths of run (MLR). They argue that the participant's improved use of prefabricated sequence of sentences increases MLR. They add that an increase in MLR is an indication of having established productions. A significant correlation of .62 between MLR and ratings of the listeners, as a finding of the present study, attests that of Towell's.

There could be another indispensable conclusion for foreign language learners and teachers. As the results of the present study showed, the strong and significant correlation between temporal fluency and proficiency score of participants clearly attest that teachers can enormously help learners to cope with disfluency phenomena; for example, by explanation of some temporal variables (pauses, repetition, and so on), conversation strategies.

References:

Chambers, F. (1997). What do we mean by fluency? *System*, *25*, 535–544.

Ejzenberg, R. (2000). The juggling act of oral fluency: A psycho-sociolinguistic metaphor. In Riggenbach, H. (Ed.), *Perspectives on fluency* (pp. 287–314). The University of Michigan Press, Michigan.

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

Freed, B. F. (1995). What makes us think that students who study abroad become fluent? In B. F. Freed (Ed.), *Second language acquisition in a study abroad context: Studies in bilingualism 9* (pp. 123-148). Amsterdam: John Benjamins.

Fulcher, G. (1996). Testing tasks: issues in task design and the group oral. *Language Testing*, *13*, 23–51

Harrell, B. (2007). *Speech-language pathologist*. Retrieved March 26, 2010, from http: // www.muncie.K12.in.us/shsweb/speech%20and%20language%20pathologist.htm#Fluency.

Hernandez, T. A. (2016). Short-term study abroad: Perspectives on speaking gains and language contact. *Applied Language Learning*, *26*, 39–64.

Huensch, A., & Thompson, A. S. (2017). Contextualizing attitudes toward pronunciation: Foreign language learners in the US. *Foreign Language Annals*, *50*, 410-432.

Huensch, A., & Tracy-Ventura, N. (2017). L2 utterance fluency development before, during, and after residence abroad: A multidimensional investigation. *The Modern Language Journal*, *101*, 275-293.

Kang, O. (2008). Ratings of L2 Oral Performance in English: Relative Impact of Rater Characteristics and Acoustic Measures of Accentedness. *Spaan Fellow Working Papers in Second or Foreign Language Assessment*, 6, 181–205

Koponen, M., Riggenbach, H. (2000). Overview: Varying perspectives on fluency. In: Riggenbach, H. (Ed.), *Perspectives on fluency* (pp. 5–24). The University of Michigan Press, Michigan.

Kormos, J. & Denes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System, 32*(2), 145–164

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning, 40,* 387– 412.

Leonard, K. R. & Shea, C. E. (2017). L2 speaking development during study abroad: fluency, accuracy, complexity, and underlying cognitive factors. *Modern Language Journal, 101*(1), 1-15.

Luoma, S. (2004). *Assessing speaking.* Cambridge: Cambridge University Press.

Mizera, G. J. (2005).*Working memory and L2 oral fluency* (Unpublished Ph.D dissertation). University of Pittsburgh, Pittsburgh.

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

Riazantseva, A. (2001). Second language proficiency and pausing. *Studies in Second Language Acquisition, 23,* 497–526.

Richards, J.C. (2006). *Communicative language teaching today.* New York: Cambridge University Press.

Richards, J.C., & Schmidt, R. (2002). *Longman dictionary of language teaching and applied linguistics* (3rd ed.). Pearson Education Limited.

Riggenbach, H. (1991). Toward an Understanding of Fluency: A Microanalysis of Nonnative Speaker Conversations. *Discourse Processes, 14*, 423-441.

Sajavaara, K. (1987). Second language speech production: Factors affecting fluency. In Dechert, H.W., Raupach, M. (Eds.), *Psycholinguistic models of production* (pp. 45–65). Ablex, Norwood.

Skehan, P. (2003). Task-based instruction. *Language Teaching, 36*, 1-14

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics, 17*, 84–119.

Ushigusa, Sh. (2008). *The Relationships between Oral Fluency and Multiword Units.* (Unpublished doctoral dissertation). Purdue University-Purdue.

Wood, D. (2006). Uses and functions of formulaic sequences in second language speech: An exploration of the foundations of fluency. *Canadian Modern Language Review, 63*, 13–33.

Wood, D. (2010). *Formulaic language and second language speech fluency*. London: Continuum International Publishing Group.