Taiwan Journal of TESOL
Vol. 17.1, 57-86, 2020 DOI: 10.30397/TJTESOL.202004_17(1).0003

SPEECH RATE AND PAUSING IN ENGLISH: COMPARING LEARNERS AT DIFFERENT LEVELS OF PROFICIENCY WITH NATIVE SPEAKERS

Lan-Fen Huang
Tomáš Gráf

ABSTRACT

In its descriptors for oral fluency the Common European Framework of Reference includes frequent references to the tempo of speech and the use of pausing. The present study aims to provide empirical evidence that these fluency phenomena exist; it also seeks to establish how they distinguish two different levels of proficiency (B2 and C1) among L2 English speakers as well as between these speakers and native speakers. The analysis is based on a large dataset of 50 native English speakers and 89 learners of English (derived from the LOCNEC corpus, and from the Czech and Taiwanese sub-corpora of LINDSEI); it compares their speech rates and the frequency and location of unfilled pauses in picture description tasks. Significant differences are found between all the groups showing that even between B2 and C1 levels the growth of proficiency is accompanied by an increase in speech rate and a decrease in the frequency of pausing, particularly within clauses and within constituents. By establishing the ranges of these values, the study sets possible targets which can be exploited for the purposes of language teaching and assessment.

Key words: learner corpus research, fluency variables, speech rates, unfilled pauses

INTRODUCTION

Despite the immense growth of learner corpus research in the context of English as a foreign language, little research so far seems to have focused on spoken learner English at different levels of proficiency. It is not hard to see why this is the case. Learner corpora, as typical sources of data for research into learner language, rarely contain reliable information about learner proficiency. They are typically compiled on the basis of the learners' first language (L1), with their second language (L2) proficiency often simply being inferred from their educational background (i.e. university students majoring in English), or at most estimated by rating a small sample of the data. For instance, ten per cent of the learners in the first eleven sub-corpora of the Louvain International Database of Spoken English Interlanguage (LINDSEI) (Gilquin, De Cock, & Granger, 2010) were assessed on the scales of the Common European Framework of Reference for Languages (CEFR) (Council of Europe, 2001).

But the various fields of applied linguistics might greatly benefit from proficiency-based learner language research, since it might result in identifying salient features of learner language at different proficiency levels and thus lead to a better understanding of the way in which proficiency develops. The present corpus-based study attempts to shed light on the fluency of the production of speech by English learners at the B2 and C1 levels using proficiency rated recordings and focusing on two salient features of fluency: the tempo of speech and pausing. These have been recognized as key components of L2 fluency (see e.g. Ahmadian (2012), Bosker, Pinget, Quené, Sanders, & De Jong (2012), Ellis & Barkhuizen (2005), Kormos & Dénes (2004), Skehan (2009)) and also figure as key fluency descriptors for spoken production in the CEFR. In order to understand better how these two components of fluency develop along with the growth of proficiency, a parallel spoken corpus of native British speakers is included in the analysis so as to establish whether significant differences exist not only between learners of different L1 backgrounds and native English speakers, which is a commonly explored question (e.g. German learners in Götz (2013) and Czech learners in Gráf (2015a), but also between learners at two adjacent proficiency levels on the CEFR scale.

Research has shown (see above) that learners typically speak more slowly than native speakers and produce more unfilled pauses. It is hypothesized that the growth of proficiency is accompanied by an increase in speech rate which becomes closer to that of native speakers and by a more native-like use of unfilled pauses (e.g. in terms of their frequency and placement within utterances). This present paper, from the angles of teaching and assessment, attempts to establish attainable targets of speech rates in monologues for learners reaching B2 or C1 and to provide more direct descriptions of the individual levels of proficiency than can be found in the available literature. The paper also suggests some pedagogical activities designed to develop fluency by increasing the speech rate and reducing the frequency and length of pauses.

When describing foreign language learners' ability to express themselves in speech it is customary to talk about their fluency and to refer to the impression they make without necessarily having in mind any objective measures of what fluency actually is. As has frequently been pointed out, the adjective *fluent* is typically used to describe L2 speech; if someone is a fluent speaker of a language, s/he is perceived to be able to produce language steadily and effortlessly. A performance of this kind leads to a positive evaluation of the speaker's linguistic abilities (Koponen & Riggenbach, 2000).

Fluency and Fluency Variables

Providing a foolproof definition and operationalising fluency for the purposes of research is, however, notoriously difficult. Fluency is typically understood as a multi-dimensional phenomenon resulting from the interplay of a variety of measures with varying degrees of salience. Opinions also vary as to what measures are included and how salient they are. For Lennon (1990), fluency is primarily a temporal phenomenon, whose key component is speech rate. In a later study, he defines fluency as 'the rapid, smooth, accurate, lucid and efficient translation of thought or communicative intention into language' (Lennon, 2000, p. 26). Similarly, Skehan (2009) understands fluency as speakers' ability to speak without interruption at a rate which most people expect. It has been argued that this ability may be linked to the use of prefabricated chunks and formulaic language (Towell, Hawkins, & Bazergui, 1996), or discourse markers (Hasselgren, 2002), whose fast retrieval and automatized production allows the speaker to gain time for formulating subsequent utterances and thus to speak with fewer interruptions.

Segalowitz (2010) points out that fluency is multidimensional but in another sense: it is closely linked to the smoothness (and fluency) of the underlying cognitive processes (cognitive fluency). These next have to be fluently converted into a stream of acoustical signals (utterance fluency) which are then perceived and impressionistically evaluated as more or less fluent by the hearer (perceived fluency). Götz (2013) offers a similar example of such a holistic approach by proposing three categories of fluency: productive, perceptive and non-verbal. Productive fluency is performance-based and made up of temporal variables, formulaic sequences, and speech management strategies. Perceptive fluency is based on the hearers' evaluation of such dimensions as accuracy, command of idiom, register, sentence structure, accent, intonation, pragmatic features, and lexical diversity. Non-verbal fluency refers to the use of paralinguistic features, such as gestures, facial expressions, body language, etc., which are, however, beyond the scope of research in most current learner corpora mainly because no videos accompany them.

What the many different approaches to describing and analysing fluency have in common is a basic understanding that fluency is primarily temporal in nature. Consequently, most fluency studies include either concrete temporal variables (typically easily quantifiable and capable of providing objective empirical evidence for the otherwise abstract nature of fluency) or features which reduce to objective terms the volume of meaningful text produced within a certain period of time. The most frequently discussed variables are the tempo of speech, the use of pauses, repair phenomena (repeats, self-correction, false starts) and formulaic language. As regards their capacity to distinguish levels of proficiency, Iwashita et al. (2008) show that the strongest predictors are the tempo of speech and the use of unfilled pauses. Interestingly, these are also the two variables underpinning oral fluency in the CEFR descriptors. A speaker at the lowest level of A1 is expected to speak 'with much pausing to search for expressions ...'; at the A2 level, s/he would use very short utterances where 'pauses, false starts and reformulation are very evident.' A speaker at the B1 level can speak longer and pause less often for grammatical formulation and the search for lexical items. When a speaker moves up to B2, pausing is expected to be much less frequent. A B2 speaker 'can produce stretches of language with a fairly even tempo ...', while a C1 speaker 'can express him/herself fluently and spontaneously ...' A speaker at the highest level, C2, 'can express him/herself spontaneously at length with a natural colloquial flow ...' (Council of Europe, 2001, pp. 28-29; 2018, p. 144).

Whilst these descriptors are intentionally vague so as to accommodate proficiency levels in different contexts and languages, it is clear that the CEFR correlates the development of fluency with the growth of proficiency and that different proficiency levels ought to be distinguishable partly by taking into account temporal variables of fluency. As De Jong (2016) suggests, fluency can be used as a diagnostic element in language assessment. The present study aims to establish whether or not quantifiable differences in speech rate and pausing exist between the two chosen levels B2 and C1 in spoken learner English.

Speech Rate

Speech rate is often considered to be one of the most robust predictors of perceived fluency (Bosker et al., 2012; Kormos & Dénes, 2004; Witton-Davies, 2015) and as such it has been extensively studied. It expresses how fast speech is delivered or, in other words, what volume of text is produced within a set period of time. However straightforward this might appear, nevertheless, measuring speech rate is far from unproblematic. First, it has to be determined what constitutes a unit of measurement, i.e. whether we are counting words, syllables or phonemes and how these are defined. Second, a decision has to be made whether to include pauses and any other components of the utterance which do not have propositional meaning (e.g. fillers, repeats, self-corrections, etc.). Third, a unit of time must be chosen.

Arguably, the most frequently used measure is the number of words per minute (e.g. Götz (2013)), which is calculated by dividing the total number of words by the length of speaking time in minutes. This is an easily imaginable measure especially for teachers and students but it has to be specified what counts as a word (e.g. is a contracted form in English to be counted as one word or two?) and whether all words are included (cf. Lennon's (1991) unpruned and pruned words, the latter of which does not include, for example, repeats and self-corrections). Another frequently used unit of measure is a syllable (e.g. syllables per second; Kendal (2013)). Here problems arise in that the syllabification of transcribed words may differ from the actual phonetic realisation. If a decision is made to exclude pauses, we speak of the articulation rate (e.g. Möhle (1984)). In the present study, as has become the norm in LINDSEI based studies, we express speech rate in words per minute (henceforth wpm).

Speech rates are known to vary widely even in native speech. Götz (2013, p. 15) reports that the speech rates of native speakers of English are shown in the literature to range between 120 and 260 wpm. In British English, it has been reported that the average speech rates of broadcasters on the radio are 150–170 wpm; of lecturers it is 125–160 wpm; in interviews, 160–210 wpm, while the highest speech rates are found in conversations, with 190–230 wpm (Tauroza & Allison, 1990, p. 102). In addition to communication contexts, other variables such as gender (Whiteside, 1996), age (Ramig, 1983) and speakers' emotions (Hausner,

1987) may affect speech rate. The nature of the task that the speaker is performing is also of importance. Cognitively, more demanding tasks usually result in a reduced speech rate (see e.g. Gráf (2015a)).

In fluency studies of learner language, speech rates have been considered a major component. In Götz's (2013) and Gráf's (2015a) studies of fluency, the comparability of these two studies is high, because the data are from two sub-corpora of LINDSEI (Gilquin et al., 2010) and its native counterpart, the Louvain Corpus of Native English Conversation (LOCNEC) (De Cock, 2004). Götz (2013, pp. 93-94), measuring speech rate manually, reported that the mean speech rates – regardless of task – in LOCNEC was 218 wpm and in the German sub-corpus of LINDSEI (LINDSEI-GE) was 160 wpm, ranging between 117 and 190 wpm. Using a more reliable method, Gráf (2015a, pp. 131-132) edited audio files, deleting all the utterances by the interviewers, and produced sound files containing learners' speech only. He recorded 203, 210 and 174 wpm in three individual tasks, which led to a mean speech rate among native speakers of 196 wpm. The speakers in the Czech component of LINDSEI (LINDSEI-CZ) spoke more slowly than the German learners, with a mean speech rate of 149 wpm (152 wpm in Task 1; 157 wpm in Task 2; and 138 wpm in Task 3).

Although the 50 German speakers in Götz's (2013) study of fluency are claimed to be advanced learners, Gráf (2015b) points out that the wide distribution of some of the observed fluency variables may also suggest a wider than expected distribution of their levels of proficiency. This is likely to be the case for the other sub-corpora of LINDSEI (Gilquin et al., 2010). The 100 learners in this study were from the Czech and Taiwanese components of LINDSEI. Their speaking proficiency levels were presented on the CEFR scales and ranged from B1 to C2 (Huang, Kubelec, Keng, & Hsu, 2018).

Some studies (e.g. Lennon (1990) and Freed (1995)) indicate that the longitudinal development of speech rates is connected with learners' proficiency. This factor is examined in the present study. Whilst a higher speech rate might indicate a higher proficiency, Munro and Derwing (2001) warn that for many speakers producing L2 speech too fast may reduce its intelligibility, especially if their speech is accented (Derwing & Munro, 1997). They also suggest that there may be an optimum speaking rate for non-native speakers which is neither too slow nor too fast.

Unfilled Pauses

Unfilled pauses are usually studied together with speech rates, since they greatly affect the measurement of speech tempo. In Biber et al.'s (1999, p. 1054) analysis of native speech, it was found that unfilled pauses occur more than 19,000 times per million words. They have been systematically categorized by their duration, function and location. For instance, Riggenbach (1991) classifies unfilled pauses by their duration. Those shorter than 0.2 seconds are called micro-pauses, which often occur in native speakers' speech without giving an impression of dysfluency. Those between 0.3 and 0.4 seconds are labelled hesitations and those between 0.5 and 3 seconds as pauses. In addition to the duration of pauses, Fillmore (1979) and Lennon (1990) argue that pauses are multi-functional. They are used for rhetorical, stylistic, physiological and planning purposes.

As regards location, Chambers (1997) distinguishes between natural pauses, occurring at structural junctions or at the end of semantic units, and unnatural ones, placed elsewhere. Tavakoli (2011) finds that pausing in the middle of clauses distinguishes learner narratives from those of NSs. In addition to the comparison between learners and native speakers, learners of lower proficiency produce more unfilled pauses within the 'Analysis of Speech Units' (see Foster et al. (2000)) than those who are of higher proficiency, as was ascertained by De Jong (2016) in a study where the proficiency of the learners had been established by means of a vocabulary test.

In interlanguage research, unfilled pauses have been shown to have a major effect on perceived fluency. Riggenbach's (1991) and Freed's (1995) studies indicate that speakers with relatively frequent pauses are judged less fluent, which corresponds to the CEFR fluency descriptors, as noted above. Cucchiarini et al. (2002) also maintain that, as regards perceived fluency ratings, it is the frequency of pauses that matters, not their duration. This is confirmed by more recent studies carried out by De Jong and her colleagues (2013; 2015) who report that pause length is not a distinguishing feature in proficiency.

In their two comparable data sets, German (Götz, 2013) and Czech learners (Gráf, 2015a) in LINDSEI, are found, unsurprisingly, to pause more frequently than native speakers. The present study, using Czech and Taiwanese proficiency-rated data, investigates not only the frequency but also the placement of unfilled pauses. The study is thus a contrastive analysis of learner English at two different levels of proficiency, produced by speakers of two typologically different L1s and it also provides a native-speaker benchmark.

Skehan and Shum (2014) compared the performance of L2 and L1 speakers of English and found that the latter spoke significantly faster and tended to pause more often at the ends of clauses than in the middle. They explain the superior performance of the native speakers as a result of their greater automaticity. Similar results were recorded by Skehan et al. (2016), who assume that if speech production follows the principles described by Levelt (1999), then end-clause pausing is conceptualizer-related and mid-clause pausing tends to be more formulator-related. This assumption, however, does not take into account the fact that during end-clause pausing both conceptualization and formulation may be taking place at the same time and the two processes may in fact be indistinguishable.

DATA AND RESEARCH METHODS

The data for the present study were extracted from the Czech (n = 50) (Gráf, 2017) and Taiwanese (n = 50) (Huang, 2014) components of the multinational corpus, LINDSEI (Gilquin et al., 2010). The speaking proficiency of the 100 learners was assessed on the scales of CEFR (Council of Europe, 2001) by two trained examiners, following the rating procedure developed by the LINDSEI team at the Catholic University of Louvain. Twenty-two jagged scores were examined again by a third rater, a senior trainer of examiners. The post hoc assessment ($\rho = .893$ on global assessment) resulted in a division into four proficiency groups: B1 (n = 9), B2 (n = 51), C1 (n = 38) and C2 (n = 2). While most of the Czech learners (n = 36) were at C1 level, most of the Taiwanese learners (n = 39) were at B2 (Huang et al., 2018). Because the numbers of B1 and C2 learners were found too low for statistical analysis, this study mainly investigates the 51 learners at the B2 level and the 38 at C1.

The learners' speech in the Czech sub-corpus, as in the other LINDSEI sub-corpora, was measured in three tasks: set topics (40,584 tokens; 42%), free discussion (42,850 tokens; 45%) and picture description (12,535 tokens; 13%) (Gráf, 2017, p. 27). In LINDSEI-TW, the three tasks accounted for 25,969 (37%), 35,450 (51%) and 8,158 (12%), respectively (Huang, 2014, p. 42). In order to lessen the effect of the task-related factors in this study (e.g. the difficulty and familiarity of the topics), only evidence from the last task was used.

The task of picture description was reasonably comparable in all of the interviews since the content was relatively controlled. Opportunities for interaction between learners and interviewers were infrequent.

Learners were asked to tell a story based on a sequence of four pictures and the interviewers for much of the time acted simply as listeners. In addition, the learners were given a little time to prepare their speech and therefore what they produced was likely to be fairly representative of their proficiency levels.

LINDSEI was complemented by its native corpus counterpart, LOCNEC (De Cock, 2004). The participants (n = 50) were university students in Britain. They were recruited to perform the three identical tasks listed above. Thus, in total, for this paper, 139 transcripts of picture description were extracted. Table 1 below lists the extracted corpus data under investigation.

Table 1

Extracted Corpus Data for Analysis

Speakers	B2	C1	Native speakers
Number of texts	51	38	50
Size (tokens)	10,693	11,195	7,181
Average tokens per text	210	295	144
Sources	LINDSEI-CZ and -TW		LOCNEC

LINDSEI-CZ = The Czech sub-corpus of the Louvain International Database of Spoken English Interlanguage
LINDSEI-TW = The Taiwanese sub-corpus of the Louvain International Database of Spoken English Interlanguage
LOCNEC = The Louvain Corpus of Native English Conversation

The extracted audio files were manually processed, taking out the beginning and closing turns and the time spent by the interviewees before they began to describe the pictures. The tokens were counted using the Wordlist tool in WordSmith 6 (Scott, 2012).

The time used by each speaker was measured by editing the audio files with Audacity (2013 members of the Audacity development team, 2013) and deleting any sections which were uttered by another person. The remaining utterances next formed the basis for calculating the total duration of time for each speaker. The speech rates were then measured in words per minute (wpm). As is customary in learner language studies so-called unpruned words (see Lennon (1991)) were used (i.e. words including repetitions, self-corrections, hesitation markers and filled pauses).

The frequency and location of unfilled pauses were analyzed. Although the pause length was not measured independently, it is, of course, reflected in the speech rates since pausing reduces the time in which words can be uttered. The investigation of unfilled pauses replicated Götz's (2013) approach. As discussed above, pause length has not been found to affect learners' perceived fluency (Cucchiarini et al., 2002) and neither does it appear to be indicative of language proficiency (De Jong & Bosker, 2013; De Jong et al., 2015); in fact, dysfluent speakers are more likely to deploy more frequent pauses than longer ones. But the present paper makes no distinction between the three pause-lengths which appear in the transcripts (one full stop refers to a pause shorter than 1 second; two full stops mean a medium pause between 1 and 3 seconds; three full stops indicate a pause longer than 3 seconds). This is because a pause of 1 second would be perceived in practice as very similar to one lasting 2 seconds. The Concord tool in WordSmith 6 (Scott, 2012) was adopted to search annotated unfilled pauses and pause rates were then calculated per hundred words (phw).

To determine the differences in the use of unfilled pauses, all the instances were further categorized on the basis of three locations: at clause boundaries (see the underlined stop in Example (1) below), within clauses (Example (2)) and within constituents (Example (3)). Pauses occurring at clause boundaries were taken as natural stops in lengthy utterances, whereas pauses within clauses and constituents were considered disruptive (Pawley & Syder, 1983). Appropriate pauses at clause boundaries gave both speakers and listeners a short break to enable them to continue and comprehend, but too many pauses within clauses and constituents (taken as hesitation pauses) may interfere with speakers' fluency and listeners' understanding.

(1) well after seeing these four pictures . I think the white lie is very important
(TW030, LINDSEI-TW)

(2) but the woman . didn't accept the answer

(TW006, LINDSEI-TW)

(3) she's a tourist so: she's she saw an . artist and she asked him (TW022, LINDSEI-TW)

The locations of unfilled pauses were analysed by the first author. The process of determining which pauses were at clause boundaries and within clauses was straightforward. However, determining the location of the unfilled pauses within constituents was somewhat more problematic. Consequently, all of the 382 instances of this phenomenon were double-checked by the second author to ensure maximum reliability; no changes were found necessary.

RESULTS AND DISCUSSION

Speech rate values were calculated for all of the learners and native speakers, which provide 51 values for the B2 level, 38 for the C1 level and 50 for the native speakers. These are discussed in the first section below, followed by a discussion of 3,197 unfilled pauses that we identified.

Speech Rates in Native and Learner Speech

The native speakers' speech rates (mean = 174; SD = 34), are illustrated in the right box-plot in Figure 1 below. The rates vary greatly, with the lowest rate at 106 and the highest at 265. These figures are close to the overall speech rates of native speakers, 120 – 260 wpm as shown by Götz (2013, p. 15). The mean speech rate of the learners of C1 level performing the picture description task was 142 wpm (SD = 20), ranging from 101 to 190 wpm. That of the B2 learners was 118 wpm (SD = 22), ranging between 73 and 178 wpm. Compared with the average rates, 106 – 265 wpm, by native speakers in LOCNEC, the B2 learners (see the left box-plot in Figure 1) spoke considerably more slowly but the C1 learners (see the middle box-plot in Figure 1) were within the average rates of the native speakers. Possible contributing factors may have been the learners' English proficiency and the very nature of picture description tasks, which entail a greater cognitive load, greater anxiety and more online planning than everyday conversation. As reported in the literature (Witton-Davies, 2012), speech rates in conversations are usually higher than in monologues.

The box-plots in Figure 1 also show a wider dispersion in LOCNEC (SD = 34, for Level B2 learners, 22, and Level C1, 20). The values in LOCNEC differ by 159 words from the slowest (106 wpm by Speaker EN013) to the fastest (265 wpm by Speaker EN002). Only four out of 50 native speakers speak more slowly than the average 128 wpm of the learners.

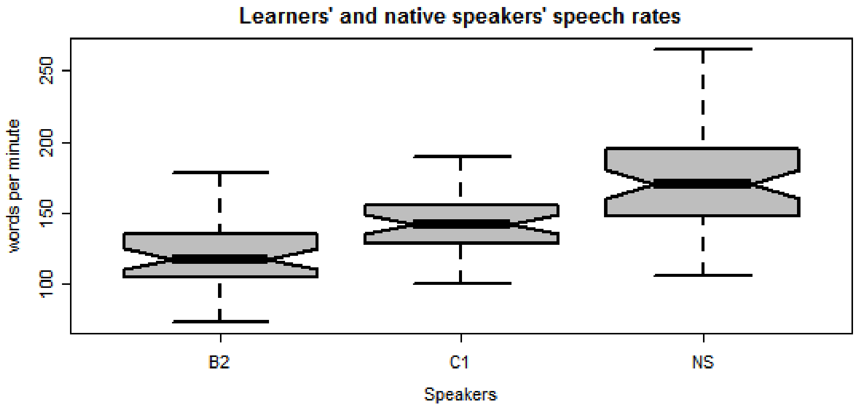

Figure 1. Box-plots of speech rates of B2 and C1 learners in LINDSEI-CZ and -TW and native speakers in LOCNEC

In relation to the two levels of English proficiency, it can be argued that the learners at Level C1 on average speak faster than those at Level B2. The speech rates of a small proportion of the C1 learners (n = 6, 15.8%) are slower than the mean of those at Level B2, while those of 13.8 per cent (n = 8) at Level B2 are faster than the mean of Level C1. A t-test resulted in $t(87) = 5.231$ ($p < .00001$). The difference in the speech rates of these two groups is statistically significant.

Compared to the 50 native speakers in LOCNEC, the 89 learners in the two proficiency groups produced significantly fewer words per minute. A one-way between-groups analysis of variance indicates that there is a statistically significant difference at the $p < .05$ level in speech rates for the three groups: F (2, 136) = 55.1, $p < .00001$. In addition to reaching statistical significance, the actual difference ($\eta 2 = .45$[1]) in mean speech

[1] Eta squared was used to calculate the effect size with ANOVA. According to Cohen (Cohen, 1988, pp. 284-287), an eta squared value of .01 is considered a small effect, .06 a

rates between the groups is quite large. Post-hoc comparisons using the Tukey HSD test indicates that the mean speech rates in the three groups are significantly different from each other. This is in line with previous studies of German (Götz, 2013) and Czech learners (Gráf, 2015a). It also confirms the hypothesis in the present paper that significant differences exist between learners and native English speakers in this regard.

The above analyses suggest that a mere investigation of speech rates may make it possible to distinguish both between the two proficiency levels and between learner and native English. It appears that a rise in proficiency between these two levels is accompanied by an increase in speech rate even at these relatively high levels of proficiency and that native speakers on average use significantly faster speech rates than advanced speakers at the C1 level.

Frequencies and Locations of Unfilled Pauses in Native and Learner Speech

In this section the frequencies and locations of unfilled pauses in native and learner speech are compared. It is hypothesized that learners pause more frequently than native speakers and the frequencies and locations of unfilled pauses in the speech of learners with a higher proficiency level (C1), compared to lower-level (B2) learners, are closer to those in the native speakers in LOCNEC. To compare the relative frequencies of unfilled pauses between LOCNEC and LINDSEI, log-likelihood values [2] were calculated in order to compare corpora of different sizes (see Table 2). Since it is unlikely for the words in a text to be normally distributed, log-likelihood is usually preferred in the analysis of texts because it does not entail any assumption about the normal distribution of words (Dunning, 1993).

The 89 learners were found to over-use unfilled pauses significantly in relation to the 50 native speakers ($G^2 > 15.13$, $p < .0001$). The same

---

medium effect and .14 a large effect.

[2] The log-likelihood values were calculated by using the LL calculator created by Rayson at Lancaster University, UK (Rayson, 2018). To have a robust indicator of how significant the differences in frequency are between two data sets, the critical values in the log-likelihood test were set at a higher value, 15.13, for the significance level of .0001; this increased the reliability (Rayson, Damon, & Brian, 2004). The critical values were 3.84, 6.63 and 10.83 for the levels of significance $p < .05$, $p < .01$ and $p < .001$ respectively. A plus or minus symbol before the log-likelihood values indicates over- or under-representation respectively in the first corpus relative to the second corpus (Rayson et al., 2004).

results were obtained by comparing native speakers with learners at B2 and C1 in turn.

Table 2

Comparison of Raw Frequencies and the Log-likelihood Values of Unfilled Pauses in LOCNEC and LINDSEI-CZ and -TW

	Speakers	Number of speakers	Size (tokens)	No of unfilled pauses	G^2
1	Learners	89	21,888	2,748	220.53
	Native speakers	50	7,181	449	($p < .0001$)
2	C1 learners	38	11,195	1,195	100.34
	Native speakers	50	7,181	449	($p < .0001$)
3	B2 learners	51	10,693	1,153	283.43
	Native speakers	50	7,181	449	($p < .0001$)
4	B2 learners	51	10,693	1,153	64.64
	C1 learners	38	11,195	1,195	($p < .0001$)

The normalized frequencies were then compared. The mean of native speakers in LOCNEC was 7.15 phw (SD = 3.94). It can be seen in Figure 2 that the means of both C1 and B2 were higher than that of the native speakers. Learners at C1, unsurprisingly, paused less often than those at B2. The mean of C1 is 10.76 phw (SD = 4.55) versus the B2 mean of 14.43 phw (SD = 4.42) ($G^2 = 64.64$, $p < .0001$). The performance of the B2 learners, in terms of the frequency of unfilled pauses, deviated more from native speakers' usage than that of the C1 learners did.

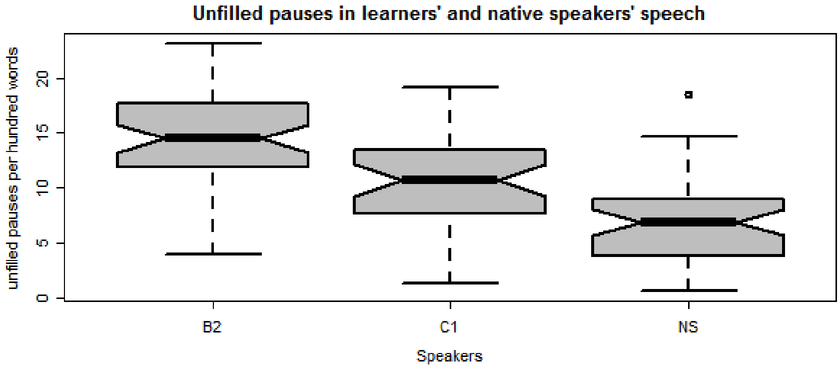

Figure 2. Box-plots of frequencies of unfilled pauses in the speech of B2 and C1 learners in LINDSEI-CZ and -TW and native speakers in LOCNEC

Table 3 below lists the mean frequency of unfilled pauses made by B2 and C1 learners in LINDSEI-CZ and -TW and native speakers in LOCNEC together with their distribution in the three locations. The mean frequency of the unfilled pauses in LOCNEC was 7.15 phw and about two thirds of them (4.64 phw) appeared at clause boundaries, whereas more than half of the unfilled pauses in B2 learner speech occurred within clauses and constituents.

Table 3

Comparison of Relative Frequencies of Unfilled Pauses in LOCNEC and LINDSEI-CZ and -TW

Speakers	**Mean frequency of UPs (phw)**	**Mean frequency of UPs at clause boundaries (phw)**	**Mean frequency of UPs within clauses (phw)**	**Mean frequency of UPs within constituents (phw)**
B2 learners	14.43 (SD = 4.42)	5.80 (SD = 2.32)	6.42 (SD = 2.18)	2.20 (SD = 1.52)
C1 learners	10.76 (SD = 4.55)	5.96 (SD = 2.62)	3.78 (SD = 2.03)	1.02 (SD = 0.93)
Native speakers	7.15 (SD = 3.94)	4.64 (SD = 2.89)	2.06 (SD = 1.97)	0.45 (SD = 0.71)

phw = per hundred words

The findings regarding the locations of unfilled pauses are illustrated in Figure 3. When the individual performances of three speaker groups were included, in some cases a considerable degree of variation became visible. The three sets of box-plots from left to right illustrate the frequencies of unfilled pauses at clause boundaries, within clauses and within constituents. The placing of unfilled pauses by B2 learners, C1 learners and native speakers is shown in the three box-plots for each location. The box-plot of unfilled pauses at clause boundaries for LOCNEC shows great variation, but there is much less variation in the native speakers' use of unfilled pauses within constituents. (Such pauses were not found in the speech of 31 of the 50 native speakers.) There were some outliers and one of the native speakers even ranged above the means of the B2 and C1 learners. This suggests that this native speaker seemed to have planning difficulties within clauses and constituents. Most native speakers in LOCNEC were well below the learners' means at either of the proficiency levels.

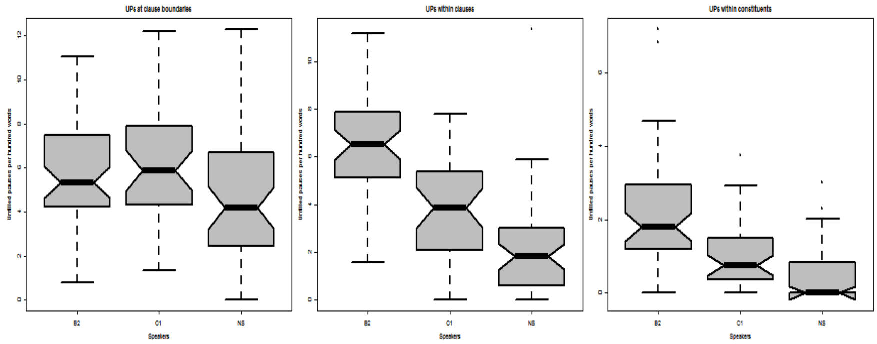

Figure 3. Box-plots of locations of unfilled pauses in the speech of B2 and C1 learners in LINDSEI-CZ and -TW and native speakers in LOCNEC

Significant differences (see Table 4) were found in the three locations under investigation between learners and native speakers, but not between B2 and C1 learners in the location of unfilled pauses at clause boundaries. Since unfilled pauses within clauses and within constituents were seen as hesitation pauses, which are not preferred by native speakers, the general expectation was that the C1 learners would use them less often than the B2 learners do. Significant differences were found in the hesitation pauses between the B2 and C1 learners. Although the C1 learners performed better than the B2, compared to native speakers, unfilled pauses remained problematic in the oral fluency of even advanced learners at C1.

Table 4

Comparison of Raw Frequencies and Log-likelihood Values of the Unfilled Pauses in Three Locations in LOCNEC and LINDSEI-CZ and -TW

	Speakers	Number of speakers	Size (tokens)	Number of unfilled pauses (UPs)	G^2	UPs at clause boundaries	G^2	UPs within clauses	G^2	UPs within consti tuents	G^2
1	**Learners**	89	21,888	2,748	220.53 (*p*<.0001)	1,291	36.72 (*p*<.0001)	1,105	161.16 (*p*<.0001)	352	73.41 (*p*<.0001)
	Native speakers	50	7,181	449		290		129		30	
2	**C1 learners**	38	11,195	1,195	100.34 (*p*<.0001)	665	28.84 (*p*<.0001)	426	62.81 (*p*<.0001)	114	21.99 (*p*<.0001)
	Native speakers	50	7,181	449		290		129		30	
3	**B2 learners**	51	10,693	1,153	283.43 (*p*<.0001)	636	31.17 (*p*<.0001)	679	223.38 (*p*<.0001)	238	111.37 (*p*<.0001)
	Native speakers	50	7,181	449		290		129		30	
4	**B2 learners**	51	10,693	1,153	64.64 (*p*<.0001)	636	0.09 (*p*>.05)	679	70.63 (*p*<.0001)	238	50.51 (*p*<.0001)
	C1 learners	38	11,195	1,195		655		426		114	

The above findings confirm those in Skehan and Shum's (2014) study and partially confirm those in a similar study done by Tavakoli (2011), in which students in an English class for the B2 level paused within clauses statistically more often than native speakers did. However, in the present study, statistically significant differences were found to lie not only in the overall frequency of pauses between B2, C1 learners and native speakers but also at the three locations between learners and native speakers. There is not enough evidence to show that C1 learners use the pauses between clauses better, but there are statistically significant differences between the two levels in using pauses within clauses and within constituents.

Fluency Performance Benchmarked Against Native Counterparts

The B2 and C1 learners' fluency performance in a monologic picture description task is summarized in Table 5 below. For the speed of speaking, the benchmark set by the native speakers in LOCNEC ranged from 106 to 265 wpm. As discussed in the previous section, the C1 learners' performance, compared to the B2 learners, is closer to that of the native speakers. As regards the frequency of pauses, learners pause more frequently than native speakers; in particular, the pause rates within constituents in B2 learners' speech (0 – 7.19 phw) are twice as high as the rates of native speakers (0 – 3 phw).

Table 5

Fluency Performance of B2 and C1 Learners and Native Counterparts

Speakers	**Speech rates (wpm)**	**Unfilled pauses (phw)**			
		Overall	At clause boundaries	Within clauses	Within constituents
B2 learners	73-178	3.97-23.11	0.79-1.04	1.58-11.18	0-7.19
C1 learners	101-190	1.33-19.15	1.33-12.18	0-7.8	0-3.76
Native speakers	106-265	0.61-18.56	0-12.28	0-11.34	0-3

phw = per hundred words
wpm = words per minute

To further evaluate the learners' fluency performance against the benchmark set by their native counterparts, the 100 learners in LINDSEI-CZ and -TW were regrouped using the native speakers' figures as cut-off points. The learners whose speech rates were slower than 106 wpm were treated as one group and those speaking within the native speakers' range formed the other group. In Table 6, it can be seen that 89 per cent of the learners whose speech rates were slower than the native speakers (below 105 wpm) were at B1 level. The proportion decreases as proficiency level rises. Ninety-seven per cent of the C1 learners and 100 per cent of the C2[3] speakers managed to speak within the range of the native speakers' benchmark. This can be valuable information in the assessment of oral ability. It is unlikely that a learner whose speech rate falls below the native range would be perceived by assessors as above B2. As regards the English Language Teaching, teachers may refer to the ranges of speech rates at CEFR levels to set achievable targets for their students.

Table 6

Distribution of Learners at Different Levels in Terms of the Speech Rate Benchmark

Speakers' levels on CEFR	**Number**	**Speech rates slower than NSs (below 105 wpm) (No. of learners/proportion)**	**Speech rates within NSs' range (106-265 wpm) (No. of learners/proportion)**
B1	9	8 (**89%**)	1 (11%)
B2	51	14 (27%)	37 (73%)
C1	38	1 (3%)	37 (**97%**)
C2	2	0 (0%)	2 (**100%**)
	100	23 (23%)	77 (77%)

CEFR = Common European Framework of Reference for Languages
wpm = words per minute

Table 7 presents the distribution of learners in terms of the benchmark of overall unfilled pause rates. The native speakers paused 0.61 to 18.56 times phw. Of the total of learners of different levels, only 11 per cent

3 Figures referring to the B1 and C1 learners are tentative, as pointed out in the section on data and research methods.

paused outside the native speakers' range. This finding seems not to be of any use; however, the distribution of learners in terms of the benchmark of unfilled pause rates within constituents is salient. In Table 8, the native speakers pause rates within constituents range from nil to 3 phw. Forty-four per cent of B1 learners and 24 per cent of B2 learners were out of this range. It can be inferred that learners at B1 and B2 still paused within constituents, which is dis-preferred, while learners at C1 and C2 were able to produce native-like pauses.

Table 7

Distribution of Learners at Different Levels in Terms of the Benchmark of Overall Unfilled Pause Rates

Speakers' levels on CEFR	**Number**	**Pause rates higher than NSs (above 18.56 phw) (Number of learners/proportion)**	**Pause rates within NSs' range (0.61-18.56 phw) (Number of learners/proportion)**
B1	9	1 (11%)	8 (89%)
B2	51	8 (16%)	43 (84%)
C1	38	2 (5%)	36 (95%)
C2	2	0 (0%)	2 (100%)
	100	11 (11%)	89 (89%)

CEFR = Common European Framework of Reference for Languages
phw = per hundred words

Table 8

Distribution of Learners at Different Levels in Terms of the Benchmark of Unfilled Pause Rates Within Constituents

Speakers' levels on CEFR	**Number**	**Pause rates within constituents higher than NSs (above 3 phw) (No. of learners/proportion)**	**Pause rates within constituents within NSs' range (0-3 phw) (No. of learners/proportion)**
B1	9	4 (**44%**)	5 (56%)
B2	51	12 (**24%**)	39 (76%)
C1	38	1 (3%)	37 (97%)
C2	2	0 (0%)	2 (100%)
	100	17 (17%)	83 (83%)

CEFR = Common European Framework of Reference for Languages
phw = per hundred words

Given the fact that both learner groups of our speakers performed the same task, i.e. reconstructing a picture story based on the same set of pictures, the more proficient speakers in our study show a greater degree of automaticity which results in the need for less time: they speak faster and need not pause as much. This may give the speakers more space for conceptualizing the message instead of formulating, which is likely to make them more fluent. Skehan et al. (2016), however, observed that speech rate and pausing habits are also affected by personal speaking styles (across different tasks but also when comparing the speakers' L1 and L2 performance) which has also been confirmed elsewhere (Derwing, Munro, Thomson, & Rossiter, 2009; De Jong et al., 2015). Such a finding cannot be verified by our study because we worked with only one picture-description task and had no recordings of our L2 speakers in their L1s.

Similarly to those of Bui et al. (2018), our results also confirm that even learners at a very advanced level of proficiency are more challenged when performing a cognitively demanding task such as the one we used in our experiment. In this, the conceptualization of the message was made harder by the simultaneous need to scan and work out the content of the pictures. Further analysis of our data may verify the frequent claim that this superior performance of the native speakers is the result of their faster access to formulaic language.

Summary

The investigation of speech rates in relation to learners' English proficiency levels indicates that there is a statistically significant relationship between learners in LINDSEI-CZ and -TW and native speakers in LOCNEC ($p < .00001$) and between two groups with differing English proficiency (B2 and C1 on CEFR) ($p < .00001$). Greater proficiency is accompanied by a faster speech rate.

Comparing the raw frequency of unfilled pauses shows that learners over-use unfilled pauses ($G^2 = 220.53$; $p < .0001$). The native speakers pause 7.15 times phw, while the learners at B2 and C1 pause much more frequently, 14.43 and 10.76 times phw, respectively. When all of the unfilled pauses were further classified on the basis of three positions—at clause boundaries, within clauses and within constituents—the difference between learners and native speakers was found significant ($p < .0001$). However, in B2 and C1 learners, significant differences were observed for the use of unfilled pauses only within clauses and within constituents, but not at clause boundaries.

The investigation of speech rates and unfilled pauses reveals that both can be used to distinguish not only native speech from learner speech but also one level of proficiency from another. Learners of both levels pause more frequently than native speakers. This confirms Pawley and Syder's (1983) claim that native speakers pause or slow down typically at or near clause boundaries and only rarely in the middle of clauses. It may be suggested that the location of pauses is an area in which learners, even advanced learners at C1, can certainly improve.

The results offer empirical data for teaching speaking within the framework of CEFR levels. From the aspect of pedagogy, this study reports reachable targets for fluency. For assessing speaking fluency, in the monologic genres, a speaker with a speech rate below 105 wpm is likely to be at the B1 level or below. Assessors can be trained with exemplars to perceive the speed of speaking. Alternatively, speech rate information may be provided by advanced media players with speech recognition technology to facilitate the appropriate CEFR rating. Another distinguishing variable in fluency, unfilled pauses within constituents, may easily be detected by trained human raters. About one fourth of B2 learners over-use it but very few C1 or more advanced learners pause unnaturally within constituents.

CONCLUSIONS

This study reveals some implications for teaching oral skills. As the investigation of speech rates shows, learners at the C1 level speak significantly faster than those at B2. Almost all (95% of) C1 speakers pause in a similar way to the native speakers in LOCNEC in terms of the frequencies and locations of unfilled pauses. B2 learners can thus be provided with fluency training so that they attain higher speech rates and a more acceptable use of pausing.

Previous studies (e.g. Arevart and Nation (1991) and Yang (2014)) argue that fluency is a trainable skill which can be improved by increasing the speech rate and reducing the frequency and length of pauses. Teachers can facilitate this learning process by devising classroom activities that give learners a chance to use what they have memorized whilst attempting to improve their pausing patterns. With lower-level learners, controlled activities may be introduced to increase learners' confidence while practicing faster speech. For example, some words, phrases and formulaic sequences in a given text may be selected for a repetition drill to train learners' pronunciation at word- or phrase-level. By improving pronunciation and imitating pausing and chunking, learners will seem more fluent (Boers, Eyckmans, Kappel, Stengers, & Demecheleer, 2006; Tavakoli, 2011).

Experiments may also be conducted using one of Schloff and Yudkin's (1991) sixty-second strategies—reading aloud an approximately 180-word text in 60 seconds. Nevertheless, the target of 180 wpm should be adjusted according to the students' proficiency levels. The length of the text, with reference to Table 5 above, may vary. For learners moving from B1 to B2, speaking at the rate of 180 wpm would be unrealistic. In addition, fluency activities can be designed to give learners enough confidence to speak English faster. The '4/3/2 technique' (Maurice, as cited in Nation, 1989), is a rehearsal-repetition activity which encourages learners to talk about the same topic faster in each of three allocated times (four, three and two minutes).

One very obvious methodological limitation in this study is that the speech under investigation represents monologic performances only. Further studies analyzing dialogic sections of the corpora are required to validate the findings presented here, since the various types of speaking require their own fluency profiles. Speaking with interlocutors is much more complex, taking into account cultural norms and turn-taking practice. Future research of this kind may be extended to other fluency variables (such as repeats, filled pauses, formulaic sequences, and discourse markers) and learner populations of other proficiency levels (such as A1, A2 and B1). In spite of these limitations, this study of two influential variables of fluency has shed some light on fluency across the proficiency levels B2 and C1. Working with a relatively large dataset of 150 speakers it has shown that speech rates and the frequency and location of unfilled pauses may not only help to distinguish learners from native speakers but also learners at different levels of proficiency and it thus provides empirical evidence for the descriptors used in CEFR.

REFERENCES

2013 members of the Audacity development team (2013). Audacity (Version 2.0.3).

Ahmadian, M. J. (2012). The effects of guided careful online planning on complexity, accuracy and fluency in intermediate EFL learners' oral production: The case of English articles. *Language Teaching Research, 16*(1), 129–149.

Arevart, S., & Nation, P. (1991). Fluency improvement in a second language. *RELC Journal, 22*(1), 84-94.

Biber, D., Finegan, E., Johansson, S., Conrad, S., & Leech, G. (1999). *Longman grammar of spoken and written English*. Essex: Pearson Education Limited.

Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research, 10*(3), 245-261.

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.

Bui, G., Skehan, P., & Wang, Z. (2018). Task condition effects on advanced-level foreign language performance. In P. A. Malovrh & A. G. Benati (Eds.), *The handbook of advanced proficiency in second language acquisition*. (pp. 219-237). Hoboken: Wiley-Blackwell.

Chambers, F. (1997). What do we mean by fluency? *System, 25*(4), 535-544.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Council of Europe (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Council of Europe (2018). *Common European framework of reference for languages: Learning, teaching, assessment companion volume with new descriptors*. Strasbourg Cedex: Council of Europe.

Cucchiarini, C., Strik, H., & Boves, L. (2002). Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. *Journal of Acoustical Society of America, 111*(6), 2862-2873.

De Cock, S. (2004). Preferred sequences of words in NS and NNS speech. *Belgian Journal of English Language and Literatures, New Series 2*, 225-246.

De Jong, N. H. (2016). Predicting pauses in L1 and L2 speech: The effects of utterance boundaries and word frequency. *International Review of Applied Linguistics in Language Teaching, 54*(2), 113-132.

De Jong, N. H., & Bosker, H. R. (2013). Choosing a threshold for silent pauses to measure second language fluency. In R. Eklund (Ed.), *Proceedings of disfluency in spontaneous speech (DiSS)* (pp. 17-20). Stockholm: Royal Institute of Technology (KTH).

De Jong, N. H., Groenhout, R., Schoonen, R., & Hulstijn, J. H. (2015). Second language fluency: Speaking style of proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics, 36*(2), 223-243.

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition, 19*(1), 1-16.
Derwing, T. M., Munro, M. J., Thomson, R. I., & Rossiter, M. J. (2009). The relationship between L1 fluency and L2 fluency development. *Studies in Second Language Acquisition, 31*(4), 533-557.
Derwing, T. M., & Munro, M. J. (2001). Modelling perceptions of the accentedness and comprehensibility of L2 speech: The role of speaking rate. *Studies in Second Language Acquisition, 23*(4), 451–468.
Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics, 19*(1), 61-74.
Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.
Fillmore, C. J. (1979). On fluency. In C. J. Fillmore, D. Kempler, & W. S-Y. Wang (Eds.), *Individual differences in language ability and language behavior* (pp. 85-101). New York: Academic Press.
Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics, 21*(3), 354-375.
Freed, B. (1995). What makes us think that students who study abroad become fluent? In B. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 123-148). Amsterdam: John Benjamin.
Götz, S. (2013). *Fluency in native and non-native English speech*. Amsterdam: John Benjamins.
Gilquin, G., De Cock, S., & Granger, S. (Eds.). (2010). *LINDSEI Louvain international database of spoken English interlanguage. Handbook and cd-rom*. Louvain-la-Neuve: Presses universitaires de Louvain.
Gráf, T. (2015a). *Accuracy and fluency in the speech of the advanced learner of English.* (Doctoral dissertation, Charles University, Prague). Retrieved from https://is.cuni.cz/webapps/zzp/detail/151663
Gráf, T. (2015b). Götz, Sandra, fluency in native and nonnative English speech. *International Journal of Learner Corpus Research, 1*(1), 178-181.
Gráf, T. (2017). The story of the learner corpus LINDSEI_CZ. *Studies in Applied Linguistics*, *8*(2), 22-35.
Hasselgren, A. (2002). Learner corpora and language testing: Smallwords as markers of learner fluency. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 143-173). Amsterdam: John Benjamins.
Hausner, M. (1987). Sprechgeschwindigkeit als eine funktion von stress: Eine fallstudie. *Language and Style, 20*, 285-311.
Huang, L.-f. (2014). Constructing the Taiwanese component of the Louvain international database of spoken English interlanguage (LINDSEI). *Taiwan Journal of TESOL, 11*(1), 31-74.
Huang, L.-f., Kubelec, S., Keng, N., & Hsu, L.-h. (2018). Evaluating CEFR rater performance through the analysis of spoken learner corpora. *Language Testing in Asia,*

8(14), 1-17.

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.

Kendall, T. (2013). *Speech rate, pause and sociolinguistic variation*. Basingstoke: Palgrave Macmillan.

Koponen, M., & Riggenbach, H. (2000). Overview: Varying perspectives on fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 5-24). Ann Arbor, MI: University of Michigan Press.

Kormos, J., & Dénes, 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*(3), 387-417.

Lennon, P. (1991). Error: Some problems of definition, identification and distinction. *Applied Linguistics, 12*(2), 180-196.

Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 25-42). Ann Arbor, MI: The University of Michigan Press.

Levelt, W. J. M. (1999). Producing spoken language: A blueprint of the speaker. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 83-122). Oxford: Oxford University Press.

Möhle, D. (1984). A comparison of the second language speech production of different native speakers. In H. W. Dechert, D. Möhle, & M. Raupach (Eds.), *Second language productions* (pp. 50-68). Tübingen: Narr.

Nation, P. (1989). Improving speaking fluency. *System, 17*(3), 377-384.

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191-226). London: Longman.

Ramig, L. A. (1983). Effects of physiological aging on speaking and reading rates. *Journal of Communication Disorders, 16*(3), 217-226.

Rayson, P. (2018). Log-likelihood and effect size calculator. Retrieved from http://ucrel.lancs.ac.uk/llwizard.html

Rayson, P., Berridge, D., & Francis, B. (2004). Extending the Cochran rule for the comparison of word frequencies between corpora. In P. G., F. C., & D. A. (Eds.), *Le poids des mots: Proceedings of the 7th international conference on statistical analysis of textual data (JADT 2004) Louvain-la-neuve, Belgium, march 10-12, 2004* (Vol. II, pp. 926 - 936). Louvain: Presses universitaires de Louvain.

Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes, 14*(4), 423-441.

Schloff, L., & Yudkin, M. (1991). *Smart speaking: Sixty-second strategies*. New York: Henry Holt.

Scott, M. (2012). Wordsmith tools (Version 6). Stroud: Lexical Analysis Software.

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

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*(4), 510-532.

Skehan, P., & Shum, S. (2014). Structure and processing condition in video-based narrative retelling. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 187-210). Amsterdam: John Benjamins.

Skehan, P., Foster, P., & Shum, S. (2016). Ladders and snakes in second language fluency. *International Review of Applied Linguistics in Language Teaching, 54*(2), 97-111.

Tauroza, S., & Allison, D. (1990). Speech rates in British English. *Applied Linguistics, 11*(1), 90-105.

Tavakoli, P. (2011). Pausing patterns: Differences between L2 learners and native speakers. *ELT Journal, 65*(1), 71-79. doi:10.1093/elt/ccq1020

Towell, R. (1987). Variability and progress in the language development of advanced learners of a foreign language. In R. Ellis (Ed.), *Second language acquisition in context*. London: Prentice Hall.

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

Whiteside, S. P. (1996). Temporal-based acoustic-phonetic patterns in read speech: Some evidence for speaker sex differences. *Journal of the International Phonetic Association, 26*(1), 23-40. doi:10.1017/S0025100300005302

Witton-Davies, G. (2012). *The variability of fluency in dialogue and monologue.* Paper presented at the Twenty-First International Symposium on English Teaching, Taipei, Taiwan.

Witton-Davies, G. (2015). Measuring oral fluency: How it can be done, and the implications for teaching and testing. In M. C. U. Department of Applied English (Ed.), *The proceedings of 2015 international conference and workshop on TEFL & applied linguistics* (pp. 311-321). Taipei: Crane Publishing.

Yang, Y. (2014). The development of speaking fluency: The 4/3/2 technique for the EFL learners in China. *International Journal of Research Studies in Language Learning, 3*(4), 55-70.

ACKNOWLEDGMENT

This work was supported by the Ministry of Science and Technology, Taiwan, under grant number MOST105-2628-H-158-001 and the Charles University project *Progres 4, Language in the shiftings of time, space, and culture*. Without these grants, the team members for this project could not have been assembled. The efforts of our project team members, Ms Shirley Hsiu-ling Chao and Ms Yu-chun Ma, are deeply appreciated. We would like to thank Professor Jer-guang Hsieh at I-Shou University for his assistance with using R. Our gratitude also goes to the LINDSEI team at the Centre for English Corpus Linguistics of the Université Catholique de Louvain, Belgium, for sharing the LOCNEC corpus. We would also like to thank the anonymous reviewers of the *Taiwan Journal of TESOL* for their constructive comments to help improve earlier versions of this paper.

CORRESPONDENCE

Lan-Fen Huang, Department of Foreign Languages, Republic of China Naval Academy, Kaohsiung, Taiwan
Email address: lanfen.huang@gmail.com

Tomáš Gráf, Department of the English Language and ELT Methodology, Charles University, Prague, Czech Republic
Email address: Tomas.Graf@ff.cuni.cz

PUBLISHING RECORD

Manuscript received: March 4, 2019; Revision received: October 30, 2019; Manuscript accepted: November 6, 2019