Article

LANGUAGE TEACHING RESEARCH

Language Teaching Research
2018, Vol. 22(1) 94–114

DOI: 10.1177/1362168816656650
journals.sagepub.com/home/ltr
SAGE

L2 fluency as influenced by content familiarity and planning: Performance, measurement, and pedagogy

Gavin Bui
Hang Seng Management College, Hong Kong

Zeping Huang
Hong Kong Baptist University, Hong Kong

Abstract
This study investigates how second language (L2) fluency is influenced by two factors: Pre-task planning and content familiarity. Planning was adopted as a between-participant variable, combined with content familiarity as a within-participant variable, in a 2 × 2 split-plot factorial design. Nineteen measures of fluency phenomena, constituting eight categories, were used. Both planning and content familiarity were found to enhance fluency, but the positive effects of planning were stronger and noticeable on a wider range of measures. The availability of planning time also helped to compensate for lack of content familiarity. Implications for pedagogy and L2 fluency measurement are discussed.

Keywords
Content familiarity, L2 fluency, pre-task planning, task-based language instruction, task-readiness

I Introduction

Research into task-based language teaching has witnessed substantial development in the past several decades as evidenced by its theoretical elaboration, methodological advancement and pedagogical connections. In particular, the construct of task planning has attracted great attention (Ellis, 2005) since the pioneering studies by Ellis (1987) and Crookes (1989). Research has shown facilitative effects of different types of

Corresponding author:
Gavin Bui, English Department, Hang Seng Management College, Siu Lek Yuen, Shatin, Hong Kong.
Email: gavinbui@hsmc.edu.hk

planning but, as argued by Bui (2014), the scope of planning is limited as to what offers 'preparedness' to learners in task completion, and the factor of content familiarity as latent planning has been largely overlooked. In addition, investigations of task conditions have typically looked at their effects on complexity, accuracy, and fluency (CAF; for a review, see Bui & Skehan, 2016, in press) and trade-offs within CAF; what has not received much attention was the differential effects within one of their performance dimensions, particularly, within the dimension of fluency. With a few exceptions (de Jong, Steinel, Florijn, Schoonen, & Hulstijn, 2012; de Jong, Groenhout, Schoonen, & Hulstijn, 2013; Kormos & Dénes, 2004), second language (L2) fluency has not been explored exclusively in task studies. Even the exceptions cited above were concerned more with the psycholinguistic nature of L2 fluency rather than from a performance perspective. Skehan (2009b) argued that fluency *per se* is a multidimensional and multifaceted construct for which subtler and deeper characterization is needed. The current study attempts to bridge the gaps in (1) the effects of content familiarity plus or minus pre-task planning on L2 oral performance and (2) the use of a finer-grained set of fluency measurements to explore these effects.

II Literature review

One of the features that characterize L2 speakers is their lack of parallel processing ability relative to native speakers (Skehan, 2009b, 2014). L2 learners typically go through serial processing in speech production, which becomes a major challenge in L2 fluency. In terms of pedagogy, there is a need to provide L2 learners with necessary facilitation in performing tasks in the L2, and planning of various sorts has been a focus of task research. Ellis (2009) distinguished between three types of planning: rehearsal, pre-task planning and within-task planning. Rehearsal gives learners opportunities for pre-task practice, and thus to improve their performance (Lynch & Maclean, 2000). The second type of planning is pre-task planning, or strategic planning, which offers learners planning time of different lengths (Mehnert, 1998) prior to actual task performance. The last type of planning pertains to within-task planning, or on-line planning. Within-task planning happens when sufficient time is available during speech production (Yuan & Ellis, 2003). Bui (2014) argues that all three forms of planning in Ellis (2009) are types of explicit planning where learners can benefit from extra preparation opportunities, whereas implicit planning types, such as content familiarity, may also get learners ready to do a task as familiarity with a certain topic is likely to reduce the need for online planning.

With either planning time or content familiarity, learners are expected to vary in the complexity, accuracy and fluency of their performances. Complexity is understood as syntactic complexity (Bui, 2014) and lexical complexity (Robinson, 2001) reflecting the learners' attempts at more advanced language. Accuracy refers to overall accuracy, indexed as error-free clauses to all clauses (Skehan & Foster, 1999) and selective formal accuracy, such as past-tense markers (Ellis, 1987). Intention to achieve better accuracy often indicates learners' conservative orientation and control over more stable elements in the interlanguage system (Skehan, 2014, p. 2). In contrast, the construct of fluency is more elusive and its operationalization differs the most among the task literature.

In layman's terms, fluency, is often used as a synonym for overall proficiency (Wood, 2010). In a narrower technical sense, fluency is viewed as an 'automatic procedural skill' in speaking (Schmidt, 1992, pp. 358–359), a process in which uttering and planning/controlling are to be executed at least partly simultaneously (Rehbein, 1987, p. 99), developed by proceduralization of knowledge of word sequences so they become prefabricated units (de Jong & Perfetti, 2011). With all this achieved, fluency manifests itself as ongoing speech without undue pausing or hesitation (Ellis & Barkhuizen, 2005).

Reviewing past work in psychology and psycholinguistics, Segalowitz (2010) proposes a distinction between three types of fluency:

- cognitive fluency, as the speaker's underlying ability to plan and deliver the speech;
- utterance fluency, as the speaker's actual performance, which should be measurable; and
- perceived fluency, as the listener or rater's estimate of one's cognitive fluency based on his or her utterance fluency.

Segalowitz links fluency to Levelt's (1999) 'blueprint' of speaking and de Bot's (1992) adaption of Levelt's (1989) version of the model to bilingual speakers, which involves two general systems: a rhetorical/semantic/syntactic system (which subsumes conceptualization and grammatical encoding), and the phonological/phonetic system (which includes morpho-phonological encoding, phonetic encoding and articulation). The first system provides a surface structure for the second one to formulate a phonological plan and execute the plan into overt speech. Segalowitz argues that except for the conceptualization stage, L2 speakers would encounter challenges at all other stages within the two systems.

Utterance fluency as a measureable construct in Segalowitz's tripartite framework is crucial to assessing L2 fluency performance, but there is no consensus regarding a measure or set of measures among researchers in the field. With diverse foci, fluency has been operationalized in a range of indices, including speed (e.g. speech rate), pausing (e.g. silence of 0.4 to 1 second), and repairs (e.g. hesitations and repetition) (for a more comprehensive account, see Freed, 2000). While complexity and accuracy are related to attention to the structure of language, fluency is more concerned with an emphasis on meaning during speaking. In an L2 context, given their limited processing capacity (Skehan, 1998), second language learners are not able to fully perform parallel processing and to simultaneously attend to fluency, complexity and accuracy. Instead, they have to make decisions on performance area(s) at which their attention is directed. Therefore, better fluency may be achieved when meaning is prioritized during L2 speaking, probably to the detriment of accuracy and complexity (for the Trade-off Hypothesis, see Skehan, 1998, 2009b; for the Cognition Hypothesis for contrasting interpretations and theorization of L2 task performance, see Robinson, 2001).

Much empirical research has pinpointed the effects of different planning conditions on L2 fluency. In general, providing L2 learners pre-task planning time of varying lengths (Mehnert, 1998) is able to increase the speed (Tavakoli & Skehan, 2005) and mean length of run (Skehan & Foster, 2005) with a reduced number of mid-clause pauses

(Foster & Skehan, 1996), all showing improvement in the temporal aspects of fluency. Such planning opportunities, on the other hand, have produced mixed results on the hesitation and repair aspects of fluency. In most analyses repairs have been taken as a fluency construct, but in others they are examined as indicators of self-correction, and measured along the dimension of accuracy (e.g. Gilabert, 2007). There is also controversy as to whether repairs are a sign of dysfluency. A certain amount of repairs in speech could 'maintain a higher rate of speech, with fewer pauses, than might otherwise be possible' but at the same time 'repair is generally accompanied by a slower speech rate' (Witton-Davies, 2010, p. 119).

As for the effects of content familiarity on L2 fluency, Bui (2014) found that topic familiarity increased the speech rate with no effect on the mean length of run. Topic familiarity also helped to cut down on repetitions but not on other repair measures such as false starts and reformulations. No other studies so far have systematically investigated content familiarity, but the work of Skehan and Foster (1999) and Bygate (2001) may shed light on how familiarity with a story structure or task types operates in an L2. They found that having prior knowledge of a storyline strongly affected L2 fluency, i.e. the more predictable the story was in the video, the more fluent learners were in narrating it. Bygate (2001) established task repetition as a very strong driving force for L2 fluency, while familiarity with the task type (not the content) was much weaker in helping learners with their speaking.

To sum up, while offering explicit planning opportunities prior to or during a task prove to be facilitative, familiarity with the content, the structure, the type or having previously performed the task also create preparedness for learners (Bui, 2014). As reviewed above, though past studies have accumulated a substantial body on the effects of planning conditions, the concept of content familiarity has largely been underexplored. On top of this, there has not been much empirical research that treats fluency as an exclusive focus area in task-based language teaching (TBLT; de Jong and her colleagues have studied L2 fluency from a purely psycholinguistic perspective and had little to say about tasks or pedagogy). There appears to be a gap between content familiarity as compared to planning and its effects on task performance, especially on L2 fluency and its possible sub-dimensions. Based on this, the following research questions are formulated to guide the current study:

- Research question 1: What are the effects of content familiarity on L2 fluency?
- Research question 2: What are the effects of pre-task planning on L2 fluency?
- Research question 3: What are the interaction effects, if any, between content familiarity and pre-task planning?
- Research question 4: What measures of L2 fluency reveal the greatest impact of different dimensions of these two task conditions?

III Methodology

1 Participants

Fifty-eight participants (21 males and 37 females) selected from 84 volunteer undergraduate students at a nursing school and a computer science department of a university

in Hong Kong participated in this study. The selection criteria included a survey of their background knowledge regarding the task topics they were expected to perform and, more critically, their proficiency levels by which groupings were made. All participants were native Cantonese speakers who had learnt English as an L2 for 12 to 15 years (13.23 years on average) at the time of the research. None of them reported staying in English-speaking countries for more than three months. Based on their self-reported English scores in the HKCEE (secondary school qualification), HKALE (official public exams for college entrance), and their C-test (see below) results, these participants were deemed to be at upper-intermediate (B2) proficiency levels according to the Common European Framework of Reference.

2 Proficiency test

The proficiency test was adapted from Dörnyei and Katona's (1992) C-test, which they found to be reliable and valid. This claim was further supported by Daller and Phelan (2006). More importantly for this research, the C-test was reported to be highly correlated with oral tasks in more recent studies (e.g. $r = .64$ in Arras, Eckes, & Grotjahn, 2002, and also in oral lexical performance in Daller & Xue, 2007). C-test was chosen as a proficiency test for two reasons: first the participants have already reported their English public exam scores, but an additional test is helpful in providing more up-to-date information regarding their current English proficiency. Second, compared to other tests, C-test was easier to administer, taking only 20–30 minutes, thus reducing fatigue during subsequent speaking tasks.

3 Speaking tasks

The two speaking tasks from Bui (2014) were used in which participants were asked to make presentations on two topics: computer viruses and natural viruses.

> General scenario: You are a specialist in the field giving a presentation to a group of university students who are neither medicine nor computer majors but are interested in the topics.
>
> Topic 1: Please describe in detail the process of a virus infecting a human body, the possible consequences, and the general procedure for dealing with a virus-infected person.
>
> Topic 2: Please describe in detail the process of a virus infecting a computer, the possible consequences, and the general procedure for dealing with a virus-infected computer.

4 Study design

Each participant was required to speak on both the computer virus and the natural virus topics. The participants' academic background and the two virus topics constituted a match and mismatch dyad. The topic of natural viruses was familiar to nursing majors, while the computer virus topic was assumed to be less familiar. The opposite assumptions applied to the computer majors. The background survey (see Section III.1) screened out candidates who reported equal familiarity in both topics, or 'reverse' direction against

this study's assumptions. For instance, a computer major who indicated better knowledge about natural viruses than computer viruses was excluded.

The 58 participants were equally divided into two cohorts for planning conditions. Half planned for 10 minutes prior to their presentations but no specific instruction was given; they could use the paper and the pen provided but they were told that the paper would be taken away when they started the presentation. The other half had to start speaking immediately after they listened to the instructions and topics. To allow meaningful analyses, the participants were instructed to produce at least 10 sentences but there was no upper limit to the length of the speech. Within each planning condition, the ordering of familiar and unfamiliar tasks was counterbalanced to eliminate possible practice effects. At the same time, each cohort constituted a nearly equal number of computer students and nursing students. Such a design aimed to ensure the robustness of the familiarity effects. Since each group was made up of similar numbers of both the computer and the nursing majors, if there is any familiarity effect, it can be assumed that it is not simply because one topic is easier than the other, because each cell performs the same two topics. Table 1 shows the grouping of the participants.

To further ensure that proficiency differences in planning conditions or in academic disciplines did not interfere with such groupings, a Univariate Analysis was carried out to test the distribution of the C-test scores, representing participants' current proficiency levels (see Table 2). Results showed that there was no significant difference between planners and non-planners, or between computer majors and nursing majors, nor was there an interaction effect between planning and major.

5 *Independent and dependent variables*

This study design has two independent variables: a between-participant variable (planning) and a within-participant variable (content familiarity). Students' academic backgrounds were equally distributed in each group to prevent possible topic bias, so their respective major *per se* did not become an independent variable.

In response to Skehan's (2009b) call for deeper and subtler characterization of the construct of fluency, and to more closely examine its multi-faceted nature, this study includes an extensive list of 19 measures as dependent variables, some of which have very rare appearances in the literature. They are differentiated into eight categories: speed, stretch, voicing, pauses at three different positions of a sentence, filled pauses, and repairs. While some measures (e.g. speech rate and mean length of run) exhibited high frequency of use in past studies, others – like the precise locations of pauses, together with the effort to distinguish between the number, the average length and the total silence time of pauses – have received less attention in fluency-relevant research methodology (Skehan, 2009b). It is important to note that the list includes composite measures that are not totally independent of the rest. For instance, phonation–time ratio is determined by the number and length of pauses.[1] The use of these composite indices aims to provide a more detailed picture of the multifaceted nature of L2 fluency and how it can be influenced by task conditions.

Table 3 describes the details of each, with a pause defined as any silence of 0.4 second or longer (following Foster & Skehan, 1996). The use of words instead of syllables per

Table 1. Study design and number of participants.

Planning (between-participant)	Familiarity (within-participant)	
	Familiar	Unfamiliar
Planners	29	29
Non-planners	29	29

Notes. The planners included 15 computer majors and 14 nursing majors, while the non-planners included 14 computer majors and 15 nursing majors.

Table 2. Proficiency difference in each group.

Source	Type III: Sum of squares	df	F	Significance
Major	5.722	1	.150	.700
Planning	8.619	1	.226	.637
Major * Planning	9.603	1	.251	.618

Table 3. Dependent variables (8 categories with 19 measures in total).

Category	Label of the measure	Description
1. Speed	Raw speech rate	Total raw words per minute. All utterances, including filled pauses (e.g. *er* and *hmm*), incomplete worlds and repairs, divided by the total duration (in minutes) of the speech.
	Pruned speech rate	Total words per minute after the deletion of filled pauses, repairs and incomplete expressions.
2. Stretch	Mean length of run	The average number of words before encountering any pause, filler or repair.
3. Voicing	Phonation time	The ratio of voicing time to the total time of utterance.
4. Mid-clause pauses	Number of mid-clause pauses	The total number of pauses in the middle of a clause per 100 words.
	Mid-clause pause length	The average length of pauses, measured in seconds, in the middle of a clause.
	Mid-clause silence total	The total silence time in seconds in the middle of a clause per 100 words.
5. Independent clause pauses	Number of independent clause pauses	The total number of pauses at the end of an independent clause per 100 words.
	Independent clause pause length	The average length of pauses, measured in seconds, at the end of an independent clause.
	Independent clause silence total	The total silence time in seconds at the end of an independent clause per 100 words.

(Continued)

Table 3. (Continued)

Category	Label of the measure	Description
6. Dependent clause pauses	Number of dependent clause pauses	The total number of pauses at the end of a dependent clause per 100 words.
	Dependent clause pause length	The average length of pauses, measured in seconds, at the end of a dependent clause.
	Dependent clause silence total	The total silence time in seconds at the end of a dependent clause per 100 words.
7. Filled pauses	Number of filled pause	The total number of filled pauses (e.g. *er, erm, hmm, eh, um, uh*) per 100 words.
	Number of pseudo filled pauses	The total number of pseudo filled pauses (e.g. *well, like, you know*) per 100 words.
8. Repairs	False start	Utterances that are abandoned before they are completed (per 100 words, same below).
	Reformulation	Repeated phrases or clauses with any modification to syntax, morphology, or word order.
	Repetition	Words, phrases or clauses that are repeated verbatim without any kind of modification.
	Replacement	Lexical items that are immediately substituted for another.

minutes is because the former has been used extensively in research of both task (e.g. Bygate, 2001; Foster & Skehan, 1996; Robinson, 2001) and oral proficiency (e.g. Freed, 2000; Lennon, 1990; Riggenbach, 1991), making future cross-study comparisons possible. Given the research foci of this paper and space limitations, results of the other two components of the CAF framework, complexity and accuracy, are not reported here. Readers can find relevant findings in Bui (2014) on whose theoretical framework this study is based.

IV Results

The statistical procedure for the split-plot design in this study with multiple dependent variables is a repeated measure multivariate analysis (MANOVA). However, a preliminary examination on the data found, not surprisingly, that some variables are highly correlated with each other. Since the application of a repeated measures MANOVA to this data set might incur multicollinearity (Meyers, Gamst, & Guarino, 2006) and hamper the accuracy of the results, an alternative procedure, a repeated measure analysis of variance (ANOVA), was adopted. The following parts report on the findings, including the means, significance level (p) and the effect size (Cohen's d), first in the text length, then in the order of the eight categories from Table 3.

Table 4 reports on the length of speeches under different familiarity levels and planning conditions. Students produced longer speeches on familiar topics than the unfamiliar ones ($p = .000$, $d = .46$), indicating that prior knowledge in a certain domain area helps with idea generation and access. Though the means show that planners produced more

Table 4. Number of words in different tasks by planners and non-planners.

Planning conditions	Familiar tasks	Unfamiliar tasks	Planning total
Non-planners (n = 29)	334.48 (147.76)	259.90 (97.95)	297.19 (122.86)
Planners (n = 29)	369.41 (177.31)	310.97 (140.51)	340.19 (158.91)
Familiarity Total (n = 58)	351.95 (162.72)	285.43 (122.77)	

Note. Standard deviation is given in parentheses.

Table 5. Fluency performance in categories 1–3 (speed, stretch, and voicing).

	Familiarity		Planning		Familiarity × planning interaction
	F	UF	PL	UnPL	
Raw speech rate	112.56 (22.04)	107.89 (24.65)	117.85 (22.05)	102.60 (22.05)	p = .000
	p = .004, d = .20		p = .008, d = .69		
Pruned speech rate	94.15 (21.60)	89.00 (25.34)	102.57 (22.12)	80.60 (19.44)	p = .005
	p = .001, d = .22		p = .000, d = 1.06		
Mean length of run	5.18 (1.49)	4.98 (1.78)	5.51 (1.96)	4.66 (1.11)	ns
	ns		p = .036, d = .53		
Phonation time	.80 (.08)	.77 (.09)	.82 (.08)	.76 (.09)	p = .000
	p = .000, d = .33		p = .000, d = 1.11		

Notes. F = familiar topic; ns = not significant; PL = planned speech; UF = unfamiliar topic; UnPL = unplanned speech. Standard deviation is given in parentheses.

words (340.19) than non-planners (297.19), the inferential statistics suggest that the difference is not significant (p = .21).

Table 5 shows the effects of content familiarity and pre-task planning on three aspects of fluency performance. Being familiar with a subject matter significantly increases the raw and the pruned speech rates, and also the phonation time, which means less silence during speaking. The same trends are found in the effects of pre-task planning, but planning produces much bigger effect sizes (more than three times larger than those by familiarity), which suggests that planning exerts a stronger influence in these areas. A notable observation from a comparison between the Raw Speech Rate and the Pruned Speech Rate is that the latter is a more sensitive measure of the influence of task characteristics and task conditions, as evidenced by the lower p value (significance) and the higher d value (effect size). In addition, familiarity does not seem to help learners produce longer stretches of words before any breakdowns or repairs, as indicated in the mean length of run. However, planning significantly increases the mean length of run with a medium effect size (d = .53). This again shows that planning is a stronger means of improving these aspects of fluency.

The interaction effects between familiarity and planning are significant in speech rates and phonation time, all pointing in the direction that the opportunity to plan prior to a task significantly narrows the difference between familiar and unfamiliar tasks.

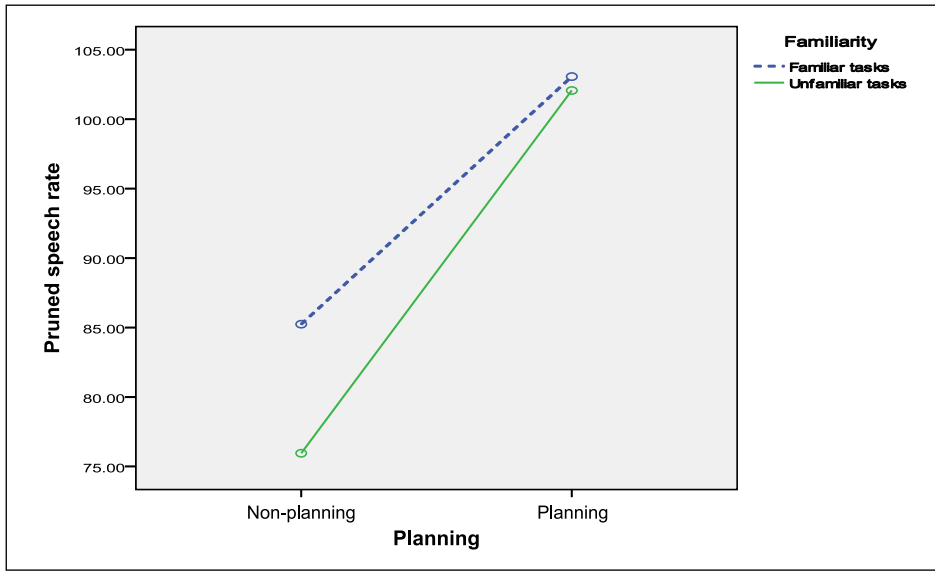

Figure 1. Familiarity × planning interaction effects in Pruned Speech Rate.

As illustrated in Figure 1 with the Pruned Speech Rate as an example, while the non-planners spoke much faster on the familiar topics than the unfamiliar ones, the planners had almost the same speech rate in both topics. Planning helped to overcome the adverse effects brought about by the unfamiliar topics.

When it comes to pausing in the middle of a clause (category 4), content familiarity and planning exhibit very similar effects. Being familiar with a topic, or being able to plan in advance, helps to reduce the number, average length and total silence of pauses in the middle of a clause, although the effect of familiarity is only approaching significance in the average length of mid-clause pauses. The effect sizes of familiarity are in general less than half of those of planning. Significant interaction effects are found in the number and the total silence of mid-clause pauses, with a similar trend found in Figure 1 that given planning time, participants are able to mitigate the difference between familiar and unfamiliar topics. Table 6 details the mid-clause pausing results.

The measure of mid-clause pauses has been widely used in the literature, but a careful distinction of end-of-clause pauses, and its further subcategories of pausing at the end of an independent and a dependent clause, from mid-clause pauses was lacking. Tables 7 and 8 report on these two sets of pauses at clausal end positions. Results in Table 7 below show that content familiarity does not have significant influence on the number, the average length, or the total silence of pauses in the three forms, at the end of an independent clause. However, planning significantly reduces pauses of these three kinds at this position, with medium to large effect sizes showing considerable effects. Interaction effects between familiarity and planning exist in the average length of pauses and the total silence here, again showing that dysfluency caused by unfamiliar topics will significantly attenuate when planning time is available.

Table 6. Fluency performance in category 4 (mid-clause pauses).

	Familiarity		Planning		Familiarity × planning interaction
	F	UF	PL	UnPL	
Number of mid-clause pauses	9.70 (5.68)	11.89 (6.91)	8.64 (5.45)	12.94 (6.28)	$p = .000$
	$p = .000$, $d = .35$		$p = .004$, $d = .73$		
Mid-clause pause length	.82 (.33)	.90 (.35)	.72 (.17)	.98 (.41)	ns
	$p = .068$, $d = .24$		$p = .000$, $d = .84$		
Mid-clause silence total	8.45 (6.36)	12.26 (13.18)	6.63 (5.09)	14.07 (11.56)	$p = .006$
	$p = .005$, $d = .37$		$p = .001$, $d = .83$		

Notes. F = familiar topic; ns = not significant; PL = planned speech; UF = unfamiliar topic; UnPL = unplanned speech. Standard deviation is given in parentheses.

Table 7. Fluency performance in category 5 (independent clause pauses).

	Familiarity		Planning		Familiarity × planning interaction
	F	UF	PL	UnPL	
Number of independent clause pauses	3.80 (1.53)	4.04 (1.54)	3.35 (1.35)	4.48 (1.30)	ns
	ns		$p = .001$, $d = .85$		
Independent clause pause length	.86 (.33)	.93 (.36)	.79 (.27)	.99 (.39)	$p = .038$
	ns		$p = .01$, $d = .60$		
Independent clause silence total	3.42 (2.06)	3.89 (2.33)	2.75 (1.53)	4.56 (2.32)	$p = .012$
	ns		$p = .000$, $d = .92$		

Notes. F = familiar topic; ns = not significant; PL = planned speech; UF = unfamiliar topic; UnPL = unplanned speech. Standard deviation is given in parentheses.

Table 8. Fluency performance in category 6 (dependent clause pauses).

	Familiarity		Planning		Familiarity × planning interaction
	F	UF	PL	UnPL	
Number of dependent clause pauses	2.87 (1.44)	2.70 (1.35)	3.08 (1.54)	2.50 (1.18)	ns
	ns		$p = .066$, $d = .42$		
Dependent clause pause length	.88 (.38)	.91 (.61)	.77 (.24)	1.03 (.64)	ns
	ns		$p = .02$, $d = .54$		
Dependent clause silence total	2.47 (1.40)	2.61 (2.21)	2.48 (1.88)	2.60 (1.76)	ns
	ns		ns		

Notes. F = familiar topic; ns = not significant; PL = planned speech; UF = unfamiliar topic; UnPL = unplanned speech. Standard deviation is given in parentheses.

Table 9. Frequency of pauses at three clausal positions.

Planning	Number of mid-clause pauses	Number of independent clause pauses	Number of dependent clause pauses
Non-planning	12.94 (6.28)	4.47 (1.51)	3.36 (1.19)
Planning	8.65 (5.46)	3.36 (1.35)	3.08 (1.54)

Note. Standard deviation is given in parentheses.

Table 10. Fluency performance in category 7 (filled pauses).

	Familiarity		Planning		Familiarity × planning interaction
	F	UF	PL	UnPL	
Number of filled pauses	10.08 (4.92)	10.73 (6.06)	9.47 (5.55)	11.35 (5.29)	ns
	ns		ns		
Number of pseudo filled pauses	.54 (.90)	.54 (.94)	.12 (.20)	.96 (1.14)	ns
	ns		p = .000, d = 1.03		

Notes. F = familiar topic; ns = not significant; PL = planned speech; UF = unfamiliar topic; UnPL = unplanned speech. Standard deviation is given in parentheses.

Results for pauses at the end of dependent clauses indicate that learners are less influenced by either content familiarity or planning in this regard. Familiarity does not have any effect for dependent clause pauses. Planning is only able to reduce the average pause length at the end of a dependent clause (p = .02, d = .54). Interestingly, planning increases the number of dependent clause pauses at this position at an almost significant level (p = .066, d = .42). No interaction effect is found.

Previous literature (e.g. Davies, 2003; Skehan, 2009a) suggests that both native and non-native speakers pause, but non-native speakers are more likely to pause in the middle of a clause while native speakers tend to pause at the end of a clause. The refined distinction between pauses at the three clausal positions in this study provides a good opportunity to verify empirically if L2 learners did pause more in the middle of rather than at the end of a clause. Table 9 shows the frequency of pauses in the middle of a clause, and at the end of both independent and dependent clauses combining data from both familiar and unfamiliar tasks. Since all these means are standardized (as frequency per 100 words) and thus become comparable, two patterns emerge. First of all, learners' pauses occur mainly in the middle of a clause. Second, when a speech is planned in advance, L2 speakers tend to cut down their mid-clause pauses rather than the end of clause pauses. Their mid-clause pauses are much more sensitive to the effect of planning than their end-of-clause pauses. In other words, planning seems to help them produce more native-like speeches in terms of the distribution of pauses.

Category 7 concerns filled pauses, on which neither content familiarity nor planning has any effects, as shown in Table 10. Filled pauses like *er* and *hmm* seem to be inherent in L2 speech. Though planning appears to significantly reduce the number of pseudo

Table 11. Fluency performance in category 8 (Repairs).

	Familiarity		Planning		Familiarity × planning interaction
	F	UF	PL	UnPL	
False start	1.45 (1.45)	1.67 (1.38)	.86 (.85)	2.27 (1.53)	ns
	ns		p = .000, d = 1.14		
Reformulation	1.48 (1.06)	1.54 (1.38)	1.08 (.84)	1.94 (1.38)	ns
	ns		p = .002, d = .75		
Repetition	3.77 (2.65)	4.78 (3.45)	2.99 (2.10)	5.56 (3.33)	ns
	p = . 001, d = .33		p = .000, d = .92		
Replacement	.96 (.78)	1.11 (1.02)	1.20 (.924)	.87 (.85)	ns
	ns		ns		

Notes. F = familiar topic; ns = not significant; PL = planned speech; UF = unfamiliar topic; UnPL = unplanned speech. Standard deviation is given in parentheses.

filled pauses such as *well, like* and *you know*, its effect should be taken with caution due to the extremely low frequency of occurrence in this measure.

Table 11 summarizes the effects of content familiarity and planning on the last category, the four repair measures. Content familiarity appears to significantly reduce the number of repetitions, although with a small effect size. It has no effect on false starts, reformulations or replacements. In contrast, planning is a strong means to decrease undesirable repairs like false starts, reformulations and repetitions (but not in replacements), with generally big effect sizes (d = .75–1.14).

The subsections above report the results of all 19 measures of fluency, including some unexpected cases. Section V disentangles the emerging patterns within these data and presents some interpretations of the findings.

V Discussion

The purpose of this study is twofold: first, to investigate the effects of content familiarity and pre-task planning on L2 fluency; second, to compare a wide range of existing and newly proposed fluency measures in one study in an attempt to contribute to the methodology of speech performance analysis. These issues are described in the following subsections, concluding with pedagogical implications.

1 Effects of content familiarity

In response to research question 1, a recap on Section IV shows that content familiarity exerts positive influence on many aspects of fluency, in particular in speech rate, phonation time, number of mid-clause pauses, mid-clause silence total, and repetition. In contrast, it does not appear as helpful in increasing the stretch of uninterrupted utterance (mean length of run), or reducing pauses and silence at the end of a clause, filled pauses and other repairs (false starts, reformulations and replacement). At this point, we can draw a few preliminary conclusions. First, content familiarity impacts the temporal

aspects of fluency, especially breakdowns in the middle of a clause. Familiarity does not help with higher automatization of speaking (mean length of run), fewer fillers or most repairs. Also, the unaffected measures of pauses at the end of an independent clause and a dependent clause indicate that L2 learners pause where they have to and providing familiar topics, though reducing the mid-clause pauses, does not change their habits of pausing at a greater grammatical juncture (end of a clause in this case).

It is then interesting to relate these results to Levelt's (1989, 1999) speaking model. Benefiting from prior knowledge (viruses in this study), learners obviously achieved a certain extent of Conceptualization and Formulation balance so as to reduce breakdowns in speaking. When familiar with the domain knowledge, L2 speakers minimized the need to gather information for idea generation. Their academic background provided them with ready-made and probably well-organized subject knowledge that could accelerate idea retrieval. Available content plus faster access to it significantly lessens the burden on attention at the Conceptualization stage. With limited processing capacity (Skehan, 1998), such ease was important for L2 learners who could allocate attention to Formulation.

Content familiarity appears to have some influence on the Formulation stage too. Retrieving lexical items and assembling them into grammatically sensible units at various levels are the first two major tasks here, especially for L2 learners at lower proficiency levels. This encoding sequence, however, may be less obvious for more advanced L2 speakers who are capable of accessing multi-word chunks and reducing the need for online processing. Participants in this study successfully decreased their mid-clause pauses and total silence while achieving higher speed, showing that they carried out less during-task planning, as compared with their unfamiliar tasks. Given prior learning in the subject (either medical or IT knowledge), participants could have developed formulaic sequences which the other group (unfamiliar) might not necessarily have. The proceduralized skill of accessing word chunks will reduce the need to pause for selection of proper words and syntactic encoding. On the other hand, content familiarity did not help to reduce the use of fillers or most types of repairs. Nor did it influence the mean length of run. This is evidence that content familiarity is much less helpful with syntactic encoding. In a familiar topic, participants still produced a similar number of repairs, especially false starts and reformulations, which involved syntactic and morphological revision to their utterances. The only type of repair that significantly diminished in the familiar topic, i.e. repetition, happened to be a verbatim duplication of a previous utterance, which was generally a lexical rather than a morpho-syntactic phenomenon. In short, content familiarity contributes more to the lexical than the syntactic encoding process.

2 *Effects of pre-task planning*

Research question 2 concerns the effects of pre-task planning. In the categories of speed, phonation time and mid-clause pauses, pre-task planning demonstrated a pattern similar to content familiarity. Nevertheless, planning appeared to be stronger (with bigger effect sizes) and more prevalent (effective in a wider range of measures). In summary, pre-task planning was capable of increasing the speed, the stretch of uninterrupted utterance, and the ratio of speaking time (phonation time), thus less unfilled pauses during speaking.

Planning reduces the number of pauses, the average length of pauses, and the total silence time in the middle of clauses and also at the ends of independent clauses. Its effects also extended to repairs where most types, including false starts, reformulations and repetitions, decreased. What appears possibly counter-intuitive but also intriguing was the impact it made at the dependent clause boundary. Planning still cut down the average length of pauses, but did not have influence on the total silence time, at this position. Though surprising at first glance, planning almost significantly increased the number of pauses at the end of a dependent clause.

The way L2 speakers utilize their planning time varies. They could sketch out a plan, find proper expressions for the topic, or even rehearse the speech in advance (Pang & Skehan, 2014). Whichever form (or combination of these forms) L2 speakers opt for, planning involves setting goals, creating ideas and organizing them into some sort of format and structure to be expressed as a speech. This process coincides with the major tasks of Conceptualization stage (Levelt, 1989) and thus could lower the workload of macro-planning during speaking. The next stage (Formulation) imposes greater pressure on L2 speakers for real-time processing and severe competition for attention allocation, as lexical, morpho-syntactic and phonological encodings are all demanding work. Native speakers are privileged with automatized skills which allow parallel processing. For L2 learners, serial processing is more likely in Formulation, and thus results in much higher demand for attentional resources. Planning appears to offer a distinct advantage to L2 learners in that not only do they have more attention saved from the Conceptualization stage for Formulation, but also they can plan their word choice and also syntax if they choose to. As such, the burden on the limited processing capacity in Formulation is greatly eased and the need for L2 learners to pause (as either filled or unfilled pauses) is therefore reduced. If a participant chooses to carry out rehearsal during planning, then the actual task performance is a repetition of the previous speech. Bygate (2001) and Wang (2014) have shown how repetition is a very powerful means of raising fluency levels in L2 learners. Thai and Boers (2016) further discovered that narrative task repetition will enhance speech rate especially in time-pressured conditions. Their interpretation of such effects is largely in line with this study. It would then appear that planning, conducted in various forms, is a helpful way to help learners with L2 fluency.

3 Interaction effects between content familiarity and pre-task planning

The contrastive element to distinguish between content familiarity and pre-task planning is the degree of external manipulation in task conditions. content familiarity prepares learners for a task in an implicit manner, with readiness inherent in the match between the task and the learner who performs the task. In contrast, pre-task planning gets learners ready by offering them extra preparation opportunities.

To answer research question 3 about the contrasts and interaction effects on content familiarity and pre-task planning, the significance values (p) suggest that planning has influence on a much wider range of fluency measures while the d values (effect sizes) indicate that when both independent variables have significant effects, the effect sizes produced by planning are usually two to three times bigger than those by content familiarity. This contrast in the magnitude of effects is due to the fact that planning allows extra

preparedness; its effects are likely to be more powerful than those of content familiarity for which the burden of during-task planning is heavier.

Interaction effects between familiarity and planning further prove that the latter is more powerful to elevate fluency performance. The magnitude of planning effects is so great that sometimes content familiarity is overridden. A case in point, as shown in Figure 1, is that given planning time, whether the presentation topic was familiar becomes unimportant for pruned speech rates, because planners achieved almost the same speed in both tasks. Though this trend is not present for every measure of fluency, the general picture has shown that planning greatly mitigates the difference between familiar and unfamiliar tasks in terms of L2 fluency, which could be of pedagogical value (to be discussed below).

It is important though to recognize that topic familiarity is a matter of degree.[2] The lesser impact of content familiarity may at least partially be due to its operationalization as the participants were somewhat familiar with and knowledgeable of the two topics in this study. Had the two topics been chosen that differed more markedly in familiarity, its effects on fluency might turn out to be stronger. Although the current study is more similar to real classrooms where teachers would not ask students to give a talk about an utterly unfamiliar topic without giving them the opportunity to look up information about it first, the gradable nature of content familiarity should be acknowledged in interpreting the results.

4 Methodology issues

For the purpose of subtler characterization of L2 fluency and better understanding of the impact of familiarity and planning, this study has employed a total of 19 measures in eight fluency aspects to provide a more comprehensive picture. Some observations in L2 fluency research methodology are listed as follows to address research question 4.

To start with, it is important to distinguish pauses at different positions, especially for L2 speakers, as their pauses between clausal boundaries are relatively rare (Davies, 2003; Tavakoli, 2011). The number of pauses, if combined together, would display a declining trend under both the planned condition and on a more familiar topic. These pauses, when dealt with separately according to their positions, reveal wider variation that might be otherwise overlooked (see Table 9). In short, mid-clause pauses are more sensitive to task manipulations, while clausal end pauses are more inherent and thus more resistant to task influence.

The second issue would be the standardization of the measures. Given different length of speeches in a task, the tallying of frequency-based measures should be made comparable through standardization. Two options are possible: a temporal length approach (occurrence per minute) and a textual length approach (occurrence per 100 words/syllable). This study argues that the textual length approach is more appropriate because the temporal method runs the risk of circular measuring in spite of its more common use in the literature (see, for example, amongst numerous other studies, Kormos & Dénes, 2004; Riggenbach, 1991). An example is that, if we are to standardize the number of pauses within a minute, the total length of these pauses becomes part of the one-minute sum and the criterion (one minute) of standardization is not independent of the variable.

In fact it is repeatedly measured. The textual length approach, in contrast, is not subject to circular calculation as it is independent of the target temporal variable.

Third, the pruned speech is a better medium for measuring the speech rate. Though both the raw speed and the pruned speed appear to significantly receive impact from various external factors, the *p* and *d* values in Table 5 suggest that the pruned speech with all fillers, repairs and incomplete words removed is more sensitive to the subtle change of task characteristics and task conditions. It is thus capable of detecting the impact of external influences with a higher level of accuracy.

Three other issues need to be addressed in future studies: one is the use of syllables per minutes instead of words per minute. Since words vary considerably in length, the calculation in syllables might more accurately capture the speech rate as it is. Though it was justified above that the use of words per minute allows future cross-study comparisons, follow-up research is encouraged to work more precisely in this regard. A second issue pertains to the collection of first language (L1) data from L2 speakers in order to use residuals rather than L2 data alone (de Jong et al., 2013). This is because one's L2 fluency may be heavily influenced by one's L1 behaviour. For example, speaking slowly in an L2 may be due to low L2 proficiency, but it may also be attributed to an individual's speaking style in their L1. That said, the effects of L2 would mostly influence between-subject comparisons, not within-subject comparisons, such as the familiar and unfamiliar topics in this study.[3] The third issue is the lack of investigation on perceived fluency. The current discussion was based on the analyses of quantitative measures, which might not be totally consistent with how people would interpret about the actual performance.

5 Pedagogical implications

Promoting greater fluency is always a major concern in L2 teaching. One of the findings of this study, consistent with task literature, was to provide pre-task time for students' strategic planning. That said, the opportunity to plan prior to speaking is not always available in daily life. An alternative is to ensure there is a good match between the task and the student. Providing students with familiar topics to talk about is one way of doing this, but factors may well go beyond content familiarity in a narrow sense. As Bui (2014) argued (see Figure 2), schematic familiarity, task type familiarity, and task repetition (as distinguished from rehearsal) may also contribute to L2 fluency as they constitute task-internal readiness. In contrast, different planning conditions are considered as task-external readiness.

This framework is useful to guide teachers' sequencing of tasks. The choice of tasks and their combinations depend largely on students' proficiency levels, but a feasible sequence of tasks may be (1) planning + familiar, (2) planning + unfamiliar, (3) non-planning + familiar, to finally (4) non-planning + unfamiliar topics. Such sequence allows students to engage in a range of learning activities from simple and easy tasks to the more difficult and demanding ones. Pre-task planning in the first two phases encourages students to gain readiness for the actual task through overt preparations. Content familiarity on the other hand provides an opportunity, without explicit rehearsal or strategic planning, for learners to achieve more fluent speech from familiarity. Therefore, the

	Macro-dimension	Micro-dimension	Sample studies
Task-readiness	• **Task-internal readiness** (implicit planning)	• **Topic familiarity** (prior subject knowledge)	Bui, 2014
		• **Schematic familiarity** (structural or procedural knowledge)	Skehan & Foster, 1999
		• **Task familiarity** (task types)	Bygate, 2001
		• **Task repetition** (content repetition without awareness of future performance.)	Bygate, 2001
	• **Task-external readiness** (explicit planning)	• **Rehearsal** (repetition with awareness of future performance)	Bei, 2013
		• **Strategic planning** (pre-task preparation)	Foster & Skehan, 1996
		• **Within-task planning** (online preparation)	Yuan & Ellis, 2003

Figure 2. A Framework of task-readiness.
Source. Bui, 2014.

latter two combinations of task conditions could be thought of as bridging classroom tasks and the real-world communication.

Bui's (2014) task-readiness framework also calls for further research into the under-explored task-internal readiness factors and the extent to which task-internal readiness can reduces the need for strategic planning and rehearsal. As planning and rehearsal will not always be available options in everyday speaking, fluent L2 speech resulted from task-internal readiness may have closer connection with what we try to prepare our learners for.

VI Conclusions

This study investigated L2 fluency as influenced by content familiarity and pre-task planning, as well as L2 fluency measurement issues. Nineteen common and newly proposed measures were employed to examine their effects on L2 fluency. Results showed that both content familiarity and pre-task planning raised temporal aspects of fluency, but pre-task planning extended its effects to the areas of repairs and pauses at the end of a clause. Planning not only had more widespread impact on fluency variables, but was more powerful in enhancing fluency performance when compared with content familiarity. This study also discusses a range of measurement issues and, based on empirical data, suggests ways to better measure L2 fluency. The findings in this study, we hope, could contribute to L2 teaching and task sequencing, as well as methodology in fluency research.

Acknowledgements

We would like to thank Peter Skehan, Frank Boers, Donovan Grose and two anonymous reviewers for their insightful comments and valuable help with this research.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The preparation of this article was supported by an RGC grant (UGC/FDS14/H01/14) that the first author received from UGC Hong Kong.

Notes

1. This was pointed out by an anonymous reviewer.
2. One anonymous reviewer pointed out this important nature of content familiarity.
3. This was an insight from one anonymous reviewer.

References

Arras, U., Eckes, T., & Grotjahn, R. (2002). C-Tests im Rahmen des 'Tests Deutsch als Fremdsprache' (TestDaF): Erste Forschungsergebnisse [C-tests as part of the 'German as a foreign language test' (TestDaF): First research results]. In: R. Grotjahn (Ed.), *Der C-Test: Theoretische Grundlangen und Praktische Anwendungen: Volume 4* [The C-test: Theoretical and practical applications: Volume 4] (pp. 175–210). Bochum: AKS-Verlag.

Bei, X.G. (2013). Effects of immediate repetition in L2 speaking task: A focused study. *English Language Teaching*, *6*, 11–19.

Bui, H.Y.G. (2014). Task readiness: Theoretical framework and empirical evidence from topic Familiarity, strategic Planning, and proficiency levels. In: P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 63–94). Amsterdam: John Benjamins.

Bui, G., & Skehan, P. (2016, in press). Complexity, fluency and accuracy. In: J. Liontas (Ed.), *TESOL encyclopedia of English language teaching*. Hoboken, NJ: Wiley-Blackwell.

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In: M. Bygate, P. Skehan, & M. Swain (Eds). *Researching pedagogical tasks: Second language learning, teaching and testing* (pp. 23–48). Harlow: Longman.

Crookes, G. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition*, *11*, 367–383.

Daller, H., & Phelan, D. (2006). The C-test and TOEIC as measurements of students' progress in intensive short courses in EFL. In: Grotjahn, R. (Ed.). *Der C-Test: Theorie, Empirie, Anwendungen* [*The C-Test: Theory, empirical research, applications*] (pp. 101–119). Frankfurt am Main: Lang.

Daller, H., & Xue, H. (2007). Lexical richness and the oral proficiency of Chinese EFL students. In: H. Daller, J. Milton, & J. Treffers-Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 150–164). Cambridge: Cambridge University Press.

Davies, A. (2003). *The native speaker: Myth and reality*. 2nd edition. Clevedon: Multilingual Matters.

de Bot, K. (1992). A bilingual production model: Levelt's 'Speaking' model adapted. *Applied Linguistics*, *13*, 1–24.

de Jong, N., & Perfetti, C.A. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language Learning*, *61*, 533–568.

de Jong, N., Groenhout, R., Schoonen, R., & Hulstijn, J.H. (2013). Second language fluency: Speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics*, *36*, 223–243.

de Jong, N., Steinel, M.P., Florijn, A., Schoonen, R., & Hulstijn, J.H. (2012). Linguistic skills and speaking fluency in a second language. *Applied Psycholinguistics*, *34*, 1–24.

Dörnyei, Z., & Katona, L. (1992). Validation of the C-test amongst Hungarian EFL learners. *Language Testing*, *9*, 187–206.

Ellis, R. (1987). Interlanguage variability in narrative discourse: Style shifting in the past tense. *Studies in Second Language Acquisition*, *9*, 1–19.

Ellis, R. (2005). Planning and task-based performance: Theory and research. In: R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 3–36). Amsterdam: John Benjamins.

Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied linguistics*, *30*, 474–509.

Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.

Foster, P., & Skehan, P. (1996). The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research*, *3*, 215–247.

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

Gilabert, R. (2007). Effects of manipulating task complexity on self-repairs during L2 oral production. *International Review of Applied Linguistics*, *45*, 214–240.

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

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, *40*, 387–417.

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

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

Lynch, T., & Maclean, J. (2000). Exploring the benefits of task repetition and recycling for classroom language learning. *Language Teaching Research*, *4*, 221–250.

Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, *20*, 83–108.

Meyers, L.S., Gamst, G., & Guarino, A.J. (2006). *Applied multivariate research: Design and interpretation*. Thousand Oaks, CA: Sage.

Pang, F., & Skehan, P. (2014). Self-reported planning behaviour and second language performance in narrative retelling. In: P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 95–127). Amsterdam: John Benjamins.

Riggenbach, H. (1991). Towards an understanding of fluency: A microanalysis of nonnative speaker conversation. *Discourse Processes*, *14*, 423–441.

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

Robinson, P. (2001). Task complexity, task difficulty and task production: Exploring interactions in a componential framework. *Applied Linguistics*, *22*, 27–57.

Schmidt, R. (1992). Psycholinguistic mechanism underlying second language fluency. *Studies in Second Language Acquisition*, *14*, 357–385.

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

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Skehan, P. (2009a). Lexical performance by native and non-native speakers on language learning tasks. In: H. Daller, D. Malvern, P. Meara, J. Milton, B. Richards, & J. Treffers-Daller (Eds.), *Vocabulary studies in first and second language acquisition: The interface between theory and application* (pp. 107–124). London: Palgrave Macmillan.

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

Skehan, P. (Ed.) (2014). *Processing perspectives on task performance*. Amsterdam: John Benjamins.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions in narrative retellings. *Language Learning*, *49*, 93–120.

Skehan, P., & Foster, P. (2005). Strategic and online planning: the influence of surprise information and task time on second language performance. In: R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 193–216). Amsterdam: Johan Benjamins.

Tavakoli, P. (2011) Pausing patterns: Differences between L2 learners and native speakers. *ELT Journal*, *65*, 71–79.

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In: R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239–273). Amsterdam: John Benjamins.

Thai, C., & Boers, F. (2016). Repeating a monologue under increasing time pressure: Effects on fluency, accuracy and complexity. *TESOL Quarterly*, *50*, 369–93.

Wang, Z. (2014). On-line time pressure manipulations: L2 speaking performance under five types of planning and repetition conditions. In: P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 27–62). Amsterdam: John Benjamins.

Witton-Davies, G. (2010). The role of repair in oral fluency. English Teachers Association ROC Conference. In *Selected papers from the Nineteenth International Symposium on English Teaching*. Taipei: Crane Publishing.

Wood, D. (2010). *Formulaic language and second language speech fluency: Background, evidence and classroom application*. London: Continuum.

Yuan, F., & Ellis, R. (2003). The effects of pre-task and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, *24*, 1–27.