

Repeating a Monologue Under Increasing Time Pressure: Effects on Fluency, Complexity, and Accuracy

CHAU THAI

*University of Danang
Danang, Vietnam*

FRANK BOERS

*Victoria University of Wellington
Wellington, New Zealand*

Studies have shown that learners' task performance improves when they have the opportunity to repeat the task. Conditions for task repetition vary, however. In the 4/3/2 activity, learners repeat a monologue under increasing time pressure. The purpose is to foster fluency, but it has been suggested in the literature that it also benefits other performance aspects, such as syntactic complexity and accuracy. The present study examines the plausibility of that suggestion. Twenty Vietnamese EFL students were asked to give the same talk three times, with or without increasing time pressure. Fluency was enhanced most markedly in the shrinking-time condition, but no significant changes regarding complexity or accuracy were attested in that condition. Although the increase in fluency was less pronounced in the constant-time condition, this increase coincided with modest gains in complexity and accuracy. The learners, especially those in the time-pressured condition, resorted to a high amount of verbatim duplication from one delivery of their narratives to the next, which explains why relatively few changes were attested in performance aspects other than fluency. The findings suggest that, if teachers wish to implement repeated-narrative activities in order to enhance output qualities beyond fluency, the 4/3/2 implementation is not the most judicious choice, and opportunities for language adjustment need to be incorporated early in the task sequence.

doi: 10.1002/tesq.232

According to Skehan's trade-off hypothesis (or limited capacity hypothesis), the simultaneous enhancement in the three speech qualities of complexity, accuracy, and fluency (CAF) is unusual.

Complexity, accuracy, and fluency “enter into competition with one another, given the limited attentional capacities of second language users” (Skehan & Foster, 1997, p. 185). For example, when a lot of attention is devoted to fluency, this will often come at the expense of attention that could be given to accuracy or complexity (Skehan, 2009, p. 512). Although there is ongoing debate about a trade-off between accuracy and complexity (e.g., Robinson, 2001, 2005), there is a substantial body of evidence to support the claim that demands on fluency may at least be at the cost of accuracy (Gass, Mackey, Alvarez-Torres, & Fernández-García, 1999; Lightbown & Spada, 1990; Robinson, 1995; Tavakoli & Skehan, 2005).

That having been said, the challenge of balancing the qualities of complexity, accuracy, and fluency in speech is likely to be influenced by the conditions of language use, which is why research into CAF frequently examines the effects of different task designs (Housen & Kuiken, 2009; Skehan, 2014). Task conditions that have been shown to affect performance include the opportunity for pretask planning (e.g., Mochizuki & Ortega, 2008); the opportunity for online planning, that is, having time to think about the way one will formulate the next part of one’s message in the course the speaking activity itself (e.g., Yuan & Ellis, 2003); and the opportunity to repeat the task (e.g., Bygate, 2001).

In this article, we investigate the interaction between opportunities for online planning and task repetition. This investigation is motivated by a practical, pedagogical concern—an evaluation of the 4/3/2 technique (Maurice, 1983) in the second or foreign language classroom. The 4/3/2 technique is intended to foster fluency development and has been advocated by, for example, Hill, Lewis, and Lewis (2000); Nation (2007); Nation and Newton (2009, pp. 153–154); and Wood (2009, 2010). In the 4/3/2 activity, language learners are asked to talk to a partner about a familiar topic for 4 minutes, repeat the talk in 3 minutes for another listener, and finally deliver the talk again in just 2 minutes for yet another listener. The increasing time pressure is meant to push the learner to convey the same content faster, more efficiently, and with fewer hesitations (in other words, more fluently). It also entails, however, that less opportunity for online planning is available to formulate the message when the monologues are repeated. There is evidence in accordance with the trade-off hypothesis that accuracy tends to be compromised in time-pressured speech (e.g., Ahmadian & Tavakoli, 2011). The question is whether the element of repetition in the 4/3/2 activity can override this tension between the exigencies of fluency and accuracy. If so, this activity can safely be utilised as a pedagogical intervention that fosters general speech quality. If not, repetition without increasing time pressure may be a judicious

alternative, especially if fluency were found to be enhanced under that condition as well. In this article we report a quasi-experimental study set up with a view to answering this question about the pedagogical merits of 4/3/2-type activities. One group of English as a foreign language (EFL) learners was asked to deliver a monologue three times under increasing time pressure, and another group was asked to deliver a monologue three times without changes to the allotted time. Before reporting the results of that study, however, we first need to discuss in more depth why the repetition of a talk, as in the 4/3/2 activity, can be expected to benefit language learners.

TASK REPETITION AND THE 4/3/2 TECHNIQUE

Task repetition, as defined by Bygate and Samuda (2005, p. 43), is the “repetition of the same or slightly altered task—whether the whole task, or parts of a task.” An early study on the benefits of task repetition is Bygate’s (1996), in which participants were required to retell the story of a cartoon video immediately after watching it and then again 10 weeks later. In another study by Bygate (2001), learners repeated narratives or interviews five times over a period of 10 weeks. More recently, Wang (2014) asked learners to twice narrate the same story they watched on video. All these experiments provide evidence of positive effects of task repetition on speech fluency, accuracy, and complexity. Additional evidence in favour of repeating speaking tasks has been furnished by Lynch and Maclean (2000) and by Ahmadian and Tavakoli (2011).

The positive effects of task repetition on fluency may reflect a process of automatisisation or proceduralisation of knowledge. The repeated use of certain utterances to perform a particular function entrenches these utterances in memory and makes them easily retrievable for future use (de Jong & Perfetti, 2011; Gatbonton & Segalowitz, 2005). That is why fluent speech tends to be formulaic, that is, made up largely of phrases that seem to be retrieved from the mental lexicon holistically, as prefabricated units (Boers, Eyckmans, Kappel, Stengers, & Demecheleer, 2006; Kuiper, 1996; Wood, 2010; Wray, 2002).

In case the task repetition concerns repetition of the same content, then the improvement in fluency may also be accounted for with reference to Levelt’s (1989) model of speech production. Speech production, according to this model, relies on the workings of three components: the conceptualiser, the formulator, and the articulator. The conceptualiser in the model plans the content of the message. If the content stays largely the same in the repeated task, then less time will need to be invested in deciding what to say as the talk is

reiterated, reducing the need for planning and thus allowing for a higher pace of delivery. The learner may then also be able to devote more processing resources to the precise wording of the (planned) content, which can enhance accuracy and complexity.

Let's now turn to the particular type of task repetition that we focus on in the present study—the 4/3/2 technique—in which increasing time pressure gradually reduces the opportunities for online planning. Although Nation and Newton (2009, p. 153) call the 4/3/2 a “well-researched activity,” to our knowledge only a handful of published studies have investigated its effectiveness. In an exploratory study, Nation (1989) conducted the activity with six learners of English as a second language (ESL) and found pronounced fluency gains between the first and the third delivery of the talks. Arevart and Nation (1991), based on a study with 20 ESL participants, confirmed this beneficial effect. It is worth noting that neither of these studies included a comparison condition in which learners repeated their talks without increasing time pressure. It is thus impossible to separate the effect of time pressure from the effect of repetition alone. It is also worth noting that the fluency gains reported in these two studies concern a comparison of the talks within the 4/3/2 sequence, not a generalised effect. A more recent study by de Jong and Perfetti (2011) suggests that incorporation of a series of 4/3/2 activities in an ESL course can produce long-term benefits for learners' fluency. It is proposed that the verbatim repetition of high-utility words and phrases across the iterations of a talk entrenches these words and phrases in the learners' second language (L2) repertoire and makes these formulas readily available to be incorporated in future speech, even if the subject matter is different. However, also in de Jong and Perfetti's investigation of fluency development, it was exclusively the 4/3/2 format of task repetition that was used, so it cannot be ruled out that repeating talks without increasing time pressure is similarly conducive to fluency development. After all, if it was the verbatim repetition of high-utility words in the 4/3/2 iterations that helped to push the participant learners' fluency, as the authors propose, then task repetition without increasing time pressure would seem to hold the same promise—unless it were found that time pressure leads speakers to resort more to verbatim repetition, which is a possibility we will revisit further below.

None of the aforementioned studies systematically investigated the effects of the 4/3/2 activity on other performance aspects besides fluency. Nation's (1989) exploratory study found some instances of improvements at the level of accuracy between the first and third deliveries of talks, but altogether the picture was mixed, with many errors preserved and some new errors emerging. Also regarding the dimension of syntactic complexity, the impressions were mixed. In a more

recent study, de Jong (2012) specifically examined the combined effects of task repetition and time pressure on the syntactic complexity of ESL learners' monologues, but she found no significant effects in that regard. She did find evidence that the increasing time pressure in the 4/3/2 format pushes speech rate significantly more than a task condition in which the same time is given for each delivery of the monologue. Like de Jong, Boers (2014) compared ESL learners' performance under conditions in which they repeated monologues either under increasing time pressure (4/3/2) or with time held constant (3/3/3). The former condition was found to push the participants' speech rate more than the latter. No compelling evidence of changes in syntactic complexity or lexical sophistication was found under either task condition. Accuracy, however, appeared to be affected differently by the two task conditions: Although a slight improvement was attested in the constant-time condition, accuracy appeared compromised in the shrinking-time condition. From a pedagogical standpoint, it suggests that the 4/3/2 technique should not be expected to generate gains beyond the fluency dimension. To reap the rewards of task repetition as far as accuracy is concerned, it may therefore be more judicious to keep the time per iteration of a talk constant (e.g., a 3/3/3 instead of 4/3/2 format).

The findings reported by Boers (2014) are based on a rather small and heterogeneous data set, however. The participants were 10 ESL learners of varying levels of proficiency and various first language (L1) backgrounds. They were asked to propose two topics they felt confident talking about and were subsequently asked to talk about one topic along the 4/3/2 sequence and about the other along a 3/3/3 sequence. It may be argued that the diversity of the monologues complicates the comparison of the two task conditions. In the present study, we therefore opted for a between-participant design, whereby all participants delivered the same type of monologue three times, either with or without increasing time pressure.

RESEARCH QUESTIONS

In this study we compared EFL learners' performance across repeated monologues with or without increasing time pressure. The aim of the study was to answer the following questions:

1. Does increasing time pressure foster fluency over and above repetition alone?
2. Does increasing time pressure compromise gains in complexity (if any) brought about by task repetition alone?

3. Does increasing time pressure compromise gains in accuracy (if any) brought about by repetition alone?

Answers to these questions can help practitioners set up the conditions under which students engage in task repetition so as to better align those conditions with the intended pedagogical outcomes of the activity.

METHOD

Participants

Twenty 10-grade EFL students at a high school in Ha Long City, Vietnam, took part in the study. They were drawn from two parallel non-English-major classes with a view to creating two equivalent groups, based on their English exam results of the previous school term. Their mean age was 15.3. All participants had started taking English as a compulsory subject from the sixth grade onwards, meaning they had been learning English for more than 4 years. They all had three 45-minute English classes per week. The intervention was incorporated in these regular classes.

Procedure

All the students were asked to give a talk on the topic of *your favourite movie*. This topic was chosen on the basis of a survey whereby the two classes of students had been presented with a list of topics and asked to indicate their degree of familiarity with each of them. Because the guidelines for implementing the 4/3/2 activity stipulate that the subject matter to be talked about should be familiar so students feel comfortable talking about it, we selected a topic—*your favourite movie*—that consistently received high familiarity ratings in the survey.

One group of students ($n = 10$) were instructed to talk about the topic first for 3 minutes, subsequently for 2 minutes, and finally for just 1 minute. We call this the shrinking-time condition. Piloting with same-profile students had revealed that it was too challenging for these learners to talk nonstop in English for 4 minutes, which is why we decided to substitute the 4/3/2 sequence with a 3/2/1 variant. The other group of students ($n = 10$) were instructed to talk about the topic for 2 minutes and then repeat their talks twice, also in a time span of 2 minutes (i.e., 2/2/2). We call this the constant-time condition. The constant-time condition is the comparison condition,

intended to help us separate the effects of increasing time pressure from the effects of task repetition alone. It is worth pointing out that the total investment of time was the same in both conditions (i.e., 6 minutes of talk time in total). Participants were given 3 minutes to prepare their talk with the aid of three questions (*What is your favourite movie? When and where did you watch it? What do you like about it?*) that would help them organise their talks.

In keeping with the guidelines for implementing the 4/3/2 technique, the students gave their talk to a new listener each time they repeated it. That way, speakers would not feel inclined to change or add content to their monologues to keep the listener entertained. This is relevant because changes in content may reduce the likelihood of verbatim repetition of words and phrases—and, as mentioned above, this verbatim repetition was identified by de Jong and Perfetti (2011) as an important factor for fluency development. It is perhaps worth mentioning in this regard that in de Jong and Perfetti's study the learners recorded their monologues on a computer instead of delivering the talk face to face. In Nation (1989) and Arevart and Nation (1991), the participants gave their talks face to face, but they took turns listening and speaking without any particular order imposed. Some speakers therefore gave their speeches after having listened to talks from their peers. This may have influenced their performances to some extent, because they had more time to think about their own talks and had the opportunity to borrow ideas and language from previous speakers. In the present experiment, strict arrangements were made so that no one would present after having listened to another participant's talk. The listeners were either classmates who were not going to perform the speaking task themselves as part of the experiment or classmates who had already finished delivering all iterations of their talk.

All the monologues (60 in total) were recorded, processed with the aid of the speech analysis programme Praat, and transcribed.

Analysis

Gauging fluency. Tavakoli and Skehan (2005) distinguish three dimensions of fluency: speech fluency, breakdown fluency, and repair fluency. Speech fluency, or speech rate, is usually measured by computing the mean number of syllables per minute. Raw speech rate is the total number of syllables per minute, but it is customary to also count only the *meaningful* syllables—that is, the syllables remaining after filled pauses, repeated syllables, and false starts are excluded. The latter speech rate has been called pruned or trimmed speech (Lennon, 1990).

Breakdown fluency and repair fluency relate to what Skehan (2014, p. 19) refers to as *flow* measures (as opposed to *speed* measures). Breakdown fluency concerns the number, length, and distribution of pauses in speech. Repair fluency concerns false starts and repetitions. Regarding breakdown fluency or the amount of conspicuous pausing, the question of how long a pause needs to be to qualify as a sign of hesitation or disfluency is a matter of some debate. For one thing, a relatively short silence in the speech of someone who habitually speaks fast may be conspicuous, while silences of the same length may be part and parcel of the smooth speech of a slow speaker. For another, the location of a silence influences the likelihood of it being perceived as a disfluency phenomenon (Tavakoli, 2011). For example, a relatively short pause between an adjective and a noun may come across as a hesitation (e.g., the speaker searching his or her mental lexicon for the appropriate noun), while a relatively long pause between two clauses or two sentences is less likely to come across as a sign of hesitation. A further complication is that one's expectations about pausing behaviour are influenced by the perceived complexity of the communicative task at hand (Derwing, Rossiter, Munro, & Thomson, 2004).

To trace the changes in fluency between our participants' iterations of their talks, we chose to examine (a) raw speech rate, that is, the mean number of syllables per minute; (b) trimmed speech rate, that is, the mean number of syllables per minute excluding false starts, repeats, and filled pauses; and (c) nonphonation/time ratio, that is, the proportion of silent pausing per minute. Taken together, these measures have the potential to capture changes in both speed and flow of oral performance.

The speech analysis software Praat (www.fon.hum.uva.nl/praat; Boersma & Weenink, 2013) was used to calculate raw speech rate per minute. De Jong and Wempe (2009) tested the reliability of Praat and conclude that "for the purpose of measuring speech rate as number of syllables per time unit... this script suffices" (p. 389). To calculate trimmed speech rate, the recordings were transcribed and any syllables, words, or phrases that were repeated, reformulated, or substituted, as well as filled pauses, were excluded. The texts were then run through the Syllable Count programme (www.syllablecount.com) to arrive at the mean number of meaningful syllables produced per minute of speech. The calculation of nonphonation/time ratio was aided by Praat again. The formula for nonphonation/time ratio is the total silent pausing time divided by the total speech time, multiplied by 100. The length of a silence required for it to be counted as pause time first has to be specified, however. We followed Wood's (2010) recommendation to set the threshold at 0.3 second.

Given the aforementioned complication that objectively measured silences need not always correspond to subjective impressions of hesitation behaviour, we will complement the results for nonphonation/time ratio with examples of conspicuously long midclause pauses, that is, pauses likely to reflect a hesitation. There were too few instances of such conspicuous midclause pauses in our samples to subject them to statistical analysis, but they can nevertheless serve to illustrate changes (or absence of change) in some participants' speech flow as they repeated their monologues.

Gauging complexity. Complexity, as defined by Ellis (2003, p. 340), is "the extent to which the language produced in performing a task is elaborate and varied." As far as syntactic complexity is concerned, many studies have examined analysis of speech (AS) unit (e.g., Norris & Ortega, 2009). Foster, Tonkyn, and Wigglesworth (2000, p. 365) define the AS unit as "a single speaker's utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clause(s) associated with either." To trace changes (if any) at the level of syntactic complexity, we thus computed the mean number of clauses per AS unit for each iteration of a talk. We adhered to Foster et al.'s guidelines, which include counting coordinated verb phrases (e.g., *It is very sad and makes me cry*) under a single AS unit, unless prosody signals otherwise.

Next to syntactic complexity, learners' output can differ markedly at the level of vocabulary. Skehan (2009) argues that, unlike native speaker discourse, in which lexical sophistication and structural complexity are often in harmony, nonnative speakers may struggle to integrate these two dimensions (potentially leading to an additional trade-off effect). Whereas most research on CAF has focused on syntax, Skehan emphasises the need to supplement such lines of research with a measure of the lexical properties of learners' output. Two types of measure can be distinguished in this respect: measures of lexical diversity and measures of lexical sophistication. Lexical diversity measures are versions of type-token counts (see, e.g., Koizumi, 2012, for an evaluation), whereas lexical sophistication measures focus on the presence of "difficult" words—operationalised as comparatively low-frequency words—in a stretch of discourse. Laufer (2005) proposes a simple measure to gauge the level of lexical sophistication of learner discourse, called *Beyond 2000*. This measure distinguishes between the 2,000 most frequent words of English (which include function words, such as articles, prepositions, and auxiliary verbs, and which by default make up the bulk of text, typically 89%–90% of native English speakers' spoken discourse; Nation, 2006; Webb & Rodgers, 2009a, 2009b) and those beyond this threshold. Despite its simplicity, the *Beyond 2000* measure is able to identify learners who have advanced beyond a

basic vocabulary (Daller, Van Hout, & Treffers-Daller, 2003). It is a measure that seems well suited for the analysis of the discourse produced in groups of intermediate-level learners such as the student participants in the present study, particularly if the focus of interest is on whether these learners substitute high-frequency words (e.g., *a film made by...*) by lower frequency words (e.g., *a film directed by...*) as they repeat a given output task. Thanks to the availability of large electronic corpora, information about word frequencies has become increasingly reliable, and this information now informs user-friendly text analysis tools that are freely available online, such as the lexical profiler on Tom Cobb's *Lexical Tutor* web site (www.lextutor.ca). To trace changes (if any) at the level of lexical sophistication between the iterations of participants' monologues, we therefore ran the trimmed speech transcripts through this lexical profiler and counted the number of words (types, not tokens) beyond the 2,000 highest frequency ones, proper nouns excepted.

Gauging accuracy. Ellis (2003, p. 339) defines accuracy as "the extent to which the language produced in performing a task conforms to target language norms." Skehan and Foster (1999, pp. 96–97) consider accuracy as a manifestation of the ability to avoid error in performance. This reflects possibly higher levels of language control but possibly also the avoidance of challenging structures that might provoke errors. Accuracy can be measured by counting non-target-like variants (or, put more simply, errors) in learners' output. Some studies have focused on particular errors (e.g., Mochizuki & Ortega, 2008; Tavakoli & Foster, 2011). A common method for gauging accuracy more generally is to determine the ratio of error-free clauses per total number of clauses (Elder & Iwashita, 2005; Foster & Skehan, 1996; Mehnert, 1998; Sangarun, 2005; Skehan & Foster, 1997; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003), and this is the method we adopted in the present study to trace changes (if any) at the level of accuracy between participants' iterations of their talks.

Verbatim duplication. In his study of the 4/3/2 activity, Boers (2014) reports a striking amount of verbatim repetition of word strings (errors included) from one delivery of a talk to the next. In many cases, more than 50% of the final versions of the monologues in his data set consisted of exact duplicates of previously uttered word strings. This inclination to resort to verbatim repetition, he argues, helps to explain the increases in speech rate, but it also leaves relatively little room for change from one delivery to the next at the level of lexical sophistication, syntactic complexity, and accuracy. As far as accuracy is concerned, verbatim repetition may even be a

double-edged sword—the repetition may not only entrench useful phrases in memory, but also carry the risk of entrenching erroneous language forms. To gauge the role of verbatim repetition in our own data, we identified all uninterrupted word strings (n-grams) in the second delivery of each monologue that were exact replicas (i.e., sequences in the same word order and with the same inflectional morphology) of the first and all the word strings in the third delivery of each monologue that were exact replicas of the first and/or second. We then calculated the proportion in percentage of words (tokens) that each repeated monologue was made up of such verbatim duplicates.

RESULTS AND DISCUSSION

Table 1 presents an overview of the changes (if any) for the measures of trimmed speech rate, clauses/AS unit, lexical sophistication, and accuracy. For the sake of simplicity we have omitted the other measures from the table, but they are included in the following sections where we report the results in more detail. In the Appendices, we provide transcripts of repeated monologues to illustrate the performance under the two conditions. To put observed changes across the three iterations of the narratives into perspective, we will first report the results of one-way ANOVAs for correlated samples. To evaluate whether the two task conditions (i.e., shrinking vs. constant time) impact the effects of task repetition differently (i.e., whether there was interaction between condition and repetition), we will report the results of two-way ANOVAs for repeated measures. (The α for statistical significance was set at < 0.5 .)

Fluency

The results for the fluency measures are consistent with previous research on the 4/3/2 technique. The mean number of syllables produced per minute of speech increased in the shrinking-time condition from 157 (*SD* 24.5) in the first delivery of the monologues over 176 (*SD* 23.0) in the second to 201 (*SD* 20.5) in the third delivery. A one-way ANOVA for correlated samples indicates this is a significant increase (F 18.51; $p < .0001$; $\eta^2 = .42$). The constant-time condition also shows an increase, from 165 (*SD* 41.5) in the first delivery over 174 (*SD* 35.5) in the second to 178 (*SD* 34.0) in the third, but this is less pronounced and falls short of significance (F 2.47; $p = .11$). A two-way ANOVA shows a significant interaction between task condition and task repetition (F 6; $p = .006$; $\eta^2 = .04$), suggesting that the

TABLE 1
Evolution of Speech Performance: Means (and Standard Deviations)

	Shrinking-time condition (3/2/1)			Significant change?	Constant-time condition (2/2/2)			Significant change?
	1st	2nd	3rd		1st	2nd	3rd	
(Trimmed) speech rate	91 (<i>SD</i> 17.5)	107 (<i>SD</i> 12)	150 (<i>SD</i> 26)	$p < .0001$	105 (<i>SD</i> 27)	116 (<i>SD</i> 22)	121 (<i>SD</i> 13)	$p = .06$
Clauses per AS unit	1.89 (<i>SD</i> 0.26)	1.66 (<i>SD</i> 0.32)	1.74 (<i>SD</i> 0.41)	$p = .63$	1.62 (<i>SD</i> 0.27)	1.81 (<i>SD</i> 0.37)	1.91 (<i>SD</i> 0.40)	$P = .003$
Lexical sophistication	3.29 (<i>SD</i> 1.57)	3.63 (<i>SD</i> 1.64)	3.78 (<i>SD</i> 1.52)	$p = .28$	3.25 (<i>SD</i> 2)	3.51 (<i>SD</i> 2.15)	3.42 (<i>SD</i> 2.19)	$p = .46$
Error-free clause ratio	0.33 (<i>SD</i> 0.12)	0.35 (<i>SD</i> 0.09)	0.33 (<i>SD</i> 0.09)	$p = .76$	0.36 (<i>SD</i> 0.15)	0.47 (<i>SD</i> 0.18)	0.53 (<i>SD</i> 0.19)	$p = .01$

increasing time pressure contributes to the increase in speech rate in the 3/2/1 activity over and above the element of task repetition alone.

The results for raw speech rate are paralleled by those for trimmed speech rate. In the shrinking-time condition, the mean number of syllables per minute increased from 91 (*SD* 17.5) over 107 (*SD* 12.0) to 150 (*SD* 26.0). This is a significant increase again (F 27.33; $p < .0001$; $\eta^2 = .64$). In the constant-time condition, the increase was less striking, from 105 (*SD* 27.0) over 116 (*SD* 22) to 121 (*SD* 13.0), and it narrowly misses significance (F 3.22; $p = .06$). A two-way ANOVA shows a significant interaction between task condition and task repetition (F 6.71; $p < .001$; $\eta^2 = .13$), suggesting again that the increasing time pressure contributes to the increase in speech rate.

The nonphonation/time ratio measure (indicating the amount of silent pausing) further corroborates the impact of time pressure. The nonphonation time ratio decreased on average by 26% between the first and third delivery under the shrinking-time condition, a significant change (F 12.3; $p < .001$; $\eta^2 = .38$). A decrease by only 5% under the constant-time condition was not significant. A two-way ANOVA shows a borderline significant interaction between task condition and task repetition for this measure as well (F 2.93; $p = .06$). Several conspicuously long within-clause pauses in students' first deliveries disappeared in their repeated versions. For example, one participant paused at the end of the utterance "Eva is a new beautiful robot from spaceship to earth to check [pause]" and then abandoned the sentence in her first delivery. She paused again the second time while apparently searching her mind for a suitable direct object: "Eva is a new, beautiful robot from spaceship sent to earth to check [pause] everything." In her final version, she duplicated the latter utterance but this time without a pause. Not all such pauses disappeared as participants repeated their monologues, however. For example, one participant paused and failed to supply the noun she was looking for in all three iterations of "It is interesting with a lot of humorous [pause]."

It is noteworthy that the shrinking-time condition led to the higher speech rates than in the constant-time condition only in the third delivery of the talks. The 3/2/1 group actually started with a comparatively slow first delivery, caught up with the 2/2/2 group in the second delivery, when both groups were given 2 minutes, and outperformed the 2/2/2 group only in the last talk, when they needed to squeeze their talks into just 1 minute. It seems likely that the first, 3-minute, window was felt to be ample time to deliver the planned content, so the students naturally spoke relatively slowly. A relatively moderate increase in speech rate was sufficient for them to convey the same information in the second delivery, but a dramatic increase was necessary if they wished to preserve much of the information in the third.

Complexity

In the shrinking-time condition, the mean ratio of clauses/AS unit fluctuates a little bit around 1.7 (see Table 1), but there is no evidence in this sample to suggest that increasing time pressure pushes syntactic complexity. In the constant-time condition, the mean ratio of clauses/AS unit rose, from 1.62 (*SD* 0.27) in the first delivery to 1.81 (*SD* 0.37) in the second and 1.91 (*SD* 0.40) in the third delivery. A one-way ANOVA shows this change to be significant (F 8.12; p = .003; η^2 = .11). However, a two-way ANOVA reveals no interaction between task condition and change in clauses/AS unit ratios (F .71; p = .5).

A good illustration of a change concerning this dimension of language use can be found in the transcripts in Appendix A, where the learner combines “After she die, he fly to the waterfall in his house” and “He use lot and lot of balloons coming from the chimney” into “After she die, he use lot of balloons coming from the chimney to fly in his house to the waterfall.” Another participant combined “The story tell about a young boy” and “He has to take care of his younger sister” into “The story is about a young boy, who has to take care of his younger sister.”

Let’s now turn to the other dimension of complexity (i.e., lexical sophistication), which we operationalise here as the number of words beyond the 2,000 most frequent lemmas of English. This measure yields no significant changes in either task condition (see Table 1). In fact, we found no examples of added Beyond 2000 words in the repeated talks produced under the shrinking-time condition. The slight increase in the mean number of such words per stretch of 100 words is a side effect of participants’ efforts at preserving as much content of their talks as possible. To do this, they needed to preserve as many information-bearing words as possible, which included the Beyond 2000 ones they had already used in the first version of their talks. We found slightly more instances of added content words in the repeated talks of the 2/2/2 group. For example, one participant produced the word *disobey* in the third delivery of his talk, while he had failed to retrieve this word in the previous deliveries. But, as we shall see further below, when participants added words, this was mostly a consequence of their adding information to their narratives.

Accuracy

The ratios of error-free clauses per total number of clauses in the trimmed speech transcripts reveal no improvement under the shrinking-time condition (see Table 1). For the 2/2/2 condition, by

contrast, there was significant improvement ($F 5.45$; $p = .014$; $\eta^2 = .15$). Mean ratios rose from .35 ($SD 0.15$) over .47 ($SD 0.18$) to .53 ($SD 0.19$). A two-way ANOVA for repeated measures reveals an interaction between the changes in accuracy and the task condition ($F 4$; $p = .027$; $\eta^2 = .05$). It therefore looks as though the potential of task repetition to push accuracy is compromised by the increasing time pressure that is the hallmark of the 4/3/2 (or 3/2/1) activity.

In fact, occasionally participants would make mistakes in the more time-pressured deliveries that they had not made in the previous deliveries. For example, one participant first produced the well-formed sentence “In the final fight between Harry and Voldemort the dark lord is defeated,” which was rendered as “He win Voldemort” in the third, time-pressured delivery of the narrative. It is worth mentioning in this context that not all learners in the shrinking-time condition dealt with the time pressure in exactly the same way. One student’s speech rate and nonphonation/time ratio hardly changed between the first and second delivery of her talk (and she did not get to the end of her narrative when time was up). This student made some improvements in accuracy, by inserting articles where these had been missing in the first delivery (e.g., “The film is about love story of . . .” → “The film is about the love story of . . .”), and consequently the ratio of error-free clauses rose from .18 in the first delivery to .4 in the second delivery of her talk. This increase in accuracy is exceptional within the 3/2/1 group, and one cannot help wondering if it might be related to the modest increase in fluency observed for this particular participant—a question we will return to below.

As already mentioned, the constant-time condition yielded more evidence of gains in accuracy from one delivery of a narrative to the next. For example, one student produced “Peeta can disguise” in his first delivery of the talk, elaborated this as “Peeta is good with disguise by painting him” in the second delivery, and corrected the form of the reflexive pronoun (“by painting himself”) in his final delivery. The same participant said “each district do one business” in his first delivery, but added the appropriate verb inflection in the subsequent versions (i.e., “each district does one business”).

Unfortunately, however, the vast majority of the errors that were made in the first delivery of the talks were preserved unmodified in subsequent talks, as we discuss in the next section.

Verbatim Duplication

The generally modest nature of changes beyond the fluency dimension (particularly in the shrinking-time condition) is better understood

when we look at the amount of verbatim repetition in the trimmed transcripts of the participants' repeated talks. In the shrinking-time condition, no less than 79.8% (*SD* 15.3) on average of the second and 90.8% (*SD* 6.5) of the third deliveries consisted of exact duplicates of word strings uttered in the preceding versions. In the constant-time condition, the means are 69.5% (*SD* 13.2) and 84.9% (*SD* 9.5), respectively. These figures are even higher than those reported in Boers (2014). It is possible that the shorter time windows in the current design (3/2/1 instead of 4/3/2) made it easier for the learners to rely on verbatim repetition. It is also possible that the 3-minute pretask planning time enabled the students not only to plan what they were going to say but also to (subvocally) rehearse how they might say it. The verbatim overlap between successive versions of talks was the greatest in the shrinking-time condition, but the difference with the constant-time condition falls short of significance (F 3.72; p = .07).

Because speech rate was enhanced also under the constant-time condition (albeit less dramatically so), participants under that condition found themselves with some extra time when they repeated their monologues, and this extra time allowed them to add content, which in turn entailed finding the necessary wording to package that additional content. Indeed, the third deliveries in the constant-time condition contained on average 28 more words (tokens) or 3.4 more clauses than the first. For example, the rather vague first account in one of the participants' talks, "She has cancer and there a man; he pay for her stay in hospital," was elaborated with information about the man and his motives in subsequent deliveries of the talk: "a young man, her manager" and "her manager, he love her." This addition of content helps to account for the slightly smaller proportion of verbatim overlap in the successive narratives in the 2/2/2 group than in the 3/2/1 group.

In any case, when up to 85% of a repeated talk is identical to a previous iteration—errors included—then it stands to reason that few changes at the levels of syntactic complexity, lexical sophistication, and accuracy are to be expected. Given that participants relied on verbatim repetition the most in the 3/2/1 condition, it is also not surprising that the fewest of these changes were attested in that data set.

Evidence of Trade-off Effects?

The above analyses suggest that task repetition with allotted time held constant yields gains (albeit sometimes marginal ones) in the speech qualities of fluency, syntactic complexity, and accuracy. This adds to the body of evidence in support of task repetition. When a

TABLE 2
Correlations Between the Gains (or Losses) on the Various Measures of Speech Performance

	Trimmed speech rate	Clauses per AS unit	Error-free clauses	Verbatim duplication
Trimmed speech rate	1	-.41*	-.41*	.27
Clauses per AS unit	-.41*	1	.12	-.21
Error-free clauses	-.41*	.12	1	-.01
Verbatim duplication	.27	-.21	-.01	1

Note. * $p < .05$.

task is repeated under increasing time pressure, however, it seems trade-offs between fluency on the one hand and complexity and accuracy on the other do not disappear entirely.

To help us evaluate whether the students who made the most pronounced progress in terms of fluency (and who would thus seem to have allocated much of their resources to this speech property) were those who made the least gain in terms of the other speech qualities (because resources were depleted), we computed Pearson's correlation coefficients between the students' ($N = 20$) gains in trimmed speech rate and changes in clauses/AS unit and error-free clause ratios (comparing the first and the final delivery of the talks). We also added the amount of verbatim duplication (between the final and first deliveries) to the analysis to assess the thesis that verbatim repetition aids fluency. For simplicity's sake, we include only one of the three measures for fluency (i.e., trimmed speech rate). The changes attested for the three fluency measures were strongly correlated, in any event (e.g., changes in trimmed and raw speech rate: $r = .86$; $p < .0001$). Because lexical sophistication and the use of subordination hardly changed between the iterations of the talks across the data set and thus showed only very minimal variance among the participants, we also omitted these two measures from the correlations.

Table 2 sums up the results, which corroborate the hypotheses: (a) increases in speech rate correlate negatively ($r = -.41$) with increases in the clause/AS units ratio (syntactic complexity; $p = .036$) as well as with increases in the error-free clause/total clauses ratio (accuracy; $p = .037$), and (b) increases in speech rate correlate positively ($r = .27$) with the amount of verbatim repetition, although this correlation falls short of significance ($p = .13$).

CONCLUSIONS, LIMITATIONS, AND IMPLICATIONS

The data we have reported here furnish no support for the suggestion that the 4/3/2 technique fosters not only fluency but also

accuracy and complexity. Although the addition of time pressure to the immediate repetition of monologues pushes learners' speech rate, it seems to dampen the opportunities for adjustments at the level of syntactic complexity and accuracy that task repetition as such can bring about. Repetition of narratives that allows for more online planning—for example, by keeping the allotted time constant for each delivery—looks more promising when the activity is intended to serve other purposes than fluency development alone. At the same time, it needs to be acknowledged that the changes in terms of complexity and accuracy in our data were not spectacular under the constant-time condition either.

We need to bear in mind, of course, that any conclusions drawn from the findings presented here are of necessity tentative. For one, the number of learners who participated in the study was small. For another, to accommodate the proficiency level of these learners, we have had to use an adaptation of the 4/3/2 activity (i.e., 3/2/1), in which the narratives were quite short. This is likely to have enabled these participants to rely on verbatim repetition more than would have been the case if we had conducted the study with more proficient students, capable of producing longer narratives. Another limitation to be acknowledged is that, unlike de Jong and Perfetti (2011), we did not investigate the effects of implementing regular 4/3/2-type (or 3/3/3-type) activities on learners' oral performance in the longer run. Our study was confined to the changes that occurred in learners' narratives within one repeated-narrative sequence. A comparison of the longer term outcomes of regular 4/3/2-type practice and regular 3/3/3-type practice may yield different findings from those reported here. Bearing these important caveats in mind, the data reported here do seem to support the limited capacity hypothesis and its associated trade-off hypothesis: As increasingly more attentional resources were required from the EFL participants to meet the demands on fluency, their ability to make adjustments to their output with respect to syntactic complexity and accuracy appeared compromised. From a pedagogical perspective, then, the results of this study suggest that the 4/3/2 technique is recommendable if the sole aim of the activity is to push fluency. However, if the objective is to foster other speech qualities as well, it would seem that task repetition without increasing time pressure is the more judicious option.

That having been said, our analysis did not reveal much language adjustment in the constant-time condition either. A likely reason why the changes from one delivery of a narrative to the next were generally limited as far as the qualities of complexity and accuracy are concerned is the immediate nature of the repetition in the procedures we have investigated—there is little or no opportunity for corrective

feedback, retrospective reflection, or language study between the deliveries of the monologues. In the absence of feedback or other opportunities for interlanguage adjustments, one can hardly expect learners to self-correct errors if they are unaware they are errors in the first place, nor expect them to use more sophisticated words if these are not yet available in their own linguistic resources. The improvements one can expect concern corrections of “slips of the tongue”—that is, of mistakes resulting from insufficient monitoring of one’s output. Indeed, the instances of corrections attested in the repeated monologues in our data virtually all concern issues about which the students had most probably received instruction, such as pluralisation and subject-verb agreement, but knowledge of which was not yet sufficiently proceduralised.

It is therefore probably sound to incorporate a stage between the first and second iteration of the talk during which students are given the incentive and the opportunity to reflect on their performance and make modifications where needed (e.g., Hawkes, 2012) or during which the teacher provides corrective feedback, so the version that gets repeated will contain more appropriate models for entrenchment in memory. Alternatively, students could be given guidance on language issues during pretask planning, before they deliver the first version of the monologue (e.g., Mochizuki & Ortega, 2008). An additional possibility is to make the speaking task more interactive so the speaker can benefit from feedback during the communicative activity itself (e.g., Ahmadian, 2012; Lynch & Maclean, 2000). There are indeed many episodes in the talks of the participants in our study when help given during the first delivery could have prevented repeated use of erroneous forms in subsequent deliveries. Examples of episodes when speakers failed to fill gaps in their resources were already mentioned in the results section. An additional example is one from the transcript in Appendix B, where the participant’s pausing behaviour in the first delivery of her talk indicated she felt uncertain about the past form of *die*, which she failed to produce (“His mother [pause] die”). This language problem could have been resolved by a helpful interlocutor. For lack of such help, this participant instead went on to substitute the form she apparently sensed was wrong by another infelicitous choice (“His mother dead,” using the adjective instead of the past verb form) in the subsequent versions of her talk.

Without incentives or opportunities for language adjustment early on in the task sequence, the immediate repetition of narratives might be likened to a double-edged sword—immediate repetition may serve the purpose of proceduralisation well enough, but there is no guarantee that what gets proceduralised in the learner’s mind is actually desirable.

ACKNOWLEDGMENTS

We would like to express our gratitude to the two anonymous reviewers for their insightful comments and helpful suggestions.

THE AUTHORS

Chau Thai is a lecturer of English at the University of Danang, in Vietnam. Her research interests are in the domain of second language acquisition, with a special focus on interventions intended to foster speaking skills. She holds a master's degree in applied linguistics from Victoria University of Wellington, in New Zealand.

Frank Boers is an associate professor at the School of Linguistics and Applied Language Studies of Victoria University of Wellington. His initial research interests were in the fields of lexicology and semantics, but he now publishes mostly on matters of instructed second language acquisition. He is co-editor of the journal *Language Teaching Research*.

REFERENCES

- 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, 129–149. doi:10.1177/1362168811425433
- Ahmadian, M. J., & Tavakoli, M. (2011). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research*, 15, 35–59. doi:10.1177/1362168810383329
- Arevart, S., & Nation, P. (1991). Fluency improvement in a second language. *REL C Journal*, 22, 84–94. doi:10.1177/003368829102200106
- Boers, F. (2014). A reappraisal of the 4/3/2 activity. *REL C Journal*, 45, 221–235. doi:10.1177/0033688214546964
- 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, 245–261. doi:10.1191/1362168806lr195oa
- Boersma, P., & Weenink, D. (2013). Praat: Doing phonetics by computer. Retrieved from <http://www.praat.org>
- Bygate, M. (1996). Effects of task repetition: Appraising the developing language of learners. In D. Willis & J. Willis (Eds.), *Challenge and change in language teaching* (pp. 136–146). London, England: Heinemann.
- Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks* (pp. 23–48). London, England: Longman.
- Bygate, M., & Samuda, V. (2005). Integrative planning through the use of task repetition. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 37–76). Amsterdam, the Netherlands: John Benjamins.
- Daller, H., Van Hout, R., & Treffers-Daller, J. (2003). Lexical richness in the spontaneous speech of bilinguals. *Applied Linguistics*, 24, 197–222. doi:10.1093/applin/24.2.197

- de Jong, N. (2012). Does time pressure help or hinder oral fluency? In N. de Jong, K. Juffermans, M. Keijzer, & L. Rasier (Eds.), *Papers of the Anéla 2012 applied linguistics conference* (pp. 43–52). Delft, the Netherlands: Eburon.
- 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. doi:10.1111/j.1467-9922.2010.00620.x
- de Jong, N., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*, 41, 385–390. doi:10.3758/BRM.41.2.385
- Derwing, T., Rossiter, M., Munro, M., & Thomson, R. (2004). Second language fluency: Judgments on different tasks. *Language Learning*, 54, 655–679. doi:10.1111/j.1467-9922.2004.00282.x
- Elder, C., & Iwashita, N. (2005). Planning for test performance: Does it make a difference? In R. Ellis (Ed.), *Planning and task performance in a second language*. Amsterdam, the Netherlands: John Benjamins.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford, England: Oxford University Press.
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18, 299–323. doi:10.1017/S0272263100015047
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21, 354–375. doi:10.1093/applin/21.3.354
- Gass, S., Mackey, A., Alvarez-Torres, M., & Fernández-García, M. (1999). The effects of task repetition on linguistic output. *Language Learning*, 49, 549–581. doi:10.1111/0023-8333.00102
- Gatbonton, E., & Segalowitz, N. (2005). Rethinking communicative language teaching: A focus on access to fluency. *Canadian Modern Language Review*, 61, 325–353.
- Hawkes, M. L. (2012). Using task repetition to direct learner attention and focus on form. *ELT Journal*, 66, 327–336. doi:10.1093/elt/ccr059
- Hill, J., Lewis, M., & Lewis, M. (2000). Classroom strategies, activities, and exercises. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp. 88–117). Hove, England: Language Teaching Publications.
- Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30, 461–478. doi:10.1093/applin/amp048
- Koizumi, R. (2012). Relationships between text length and lexical diversity measures: Can we use short texts of less than 100 tokens? *Vocabulary Learning and Instruction*, 1, 60–69. doi:10.7820/vli.v01.1.koizumi
- Kuiper, K. (1996). *Smooth talkers: The linguistic performance of auctioneers and sportscasters*. Mahwah, NJ: Lawrence Erlbaum.
- Laufer, B. (2005). Lexical frequency profiles: From Monte Carlo to the real world: A response to Meara. *Applied Linguistics*, 26, 582–588. doi:10.1093/applin/ami029
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40, 387–417. doi:10.1111/j.1467-1770.1990.tb00669.x
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Lightbown, P., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching. *Studies in Second Language Acquisition*, 12, 429–448. doi:10.1017/S0272263100009517

- Lynch, T., & Maclean, J. (2000). Exploring the benefits of task repetition and recycling for classroom language learning. *Language Teaching Research*, 4, 221–250. doi:10.1177/136216880000400303
- Maurice, K. (1983). The fluency workshop. *TESOL Newsletter*, 17, 429.
- Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, 20, 83–108.
- Mochizuki, N., & Ortega, L. (2008). Balancing communication and grammar in beginning-level foreign language classrooms: A study of guided planning and relativization. *Language Teaching Research*, 12, 11–37. doi:10.1177/1362168807084492
- Nation, I. S. P. (1989). Improving speaking fluency. *System*, 17, 377–384. doi:10.1016/0346-251X(89)90010-9
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63, 59–82.
- Nation, I. S. P. (2007). The four strands. *Innovation in Language Learning and Teaching*, 1, 2–13. doi:10.2167/illt039.0
- Nation, I. S. P., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. New York, NY: Routledge.
- Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30, 555–578. doi:10.1093/applin/amp044
- Robinson, P. (1995). Task complexity and second language narrative discourse. *Language Learning*, 45, 99–140. doi:10.1111/j.1467-1770.1995.tb00964.x
- Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22, 27–57. doi:10.1093/applin/22.1.27
- Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics in Language Teaching*, 43, 1–32.
- Sangarun, J. (2005). The effects of focusing on meaning and form in strategic planning. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 111–141). Philadelphia, PA: John Benjamins.
- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30, 510–532. doi:10.1093/applin/amp047
- Skehan, P. (2014). The context for researching a processing perspective on task performance. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 1–26). Philadelphia, PA: John Benjamins.
- Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, 1, 185–211. doi:10.1177/136216889700100302
- Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, 49, 93–120. doi:10.1111/1467-9922.00071
- Tavakoli, P. (2011). Pausing patterns: Differences between L2 learners and native speakers. *ELT Journal*, 65, 71–79. doi:10.1093/elt/ccq020
- Tavakoli, P., & Foster, P. (2011). Task design and second language performance: The effect of narrative type on learner output: Task design and L2 performance. *Language Learning*, 61, 37–72. doi:10.1111/j.1467-9922.2011.00642.x
- 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, the Netherlands: John Benjamins.

- 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). Philadelphia, PA: John Benjamins.
- Webb, S., & Rodgers, M. P. (2009a). The lexical coverage of movies. *Applied Linguistics*, 30, 407–427. doi:10.1093/applin/amp010
- Webb, S., & Rodgers, M. P. (2009b). Vocabulary demands of television programs. *Language Learning*, 59, 335–366. doi:10.1111/j.1467-9922.2009.00509.x
- Wood, D. (2009). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. *Canadian Journal of Applied Linguistics*, 12, 39–57.
- Wood, D. (2010). *Formulaic language and second language speech fluency: Background, evidence and classroom applications*. New York, NY: Continuum.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge, England: Cambridge University Press.
- Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, 24, 1–27. doi:10.1093/applin/24.1.1

APPENDIX A

EXAMPLE OF REPEATED MONOLOGUE (TRIMMED SPEECH) IN THE SHRINKING-TIME CONDITION (3/2/1)

First Delivery

I like to watch cartoon in my free time. My favourite cartoon is Up. It a story about an old man. He met his wife when they both little kid. They are together until they get old and she die. They always want to go to a great waterfall in South America but they couldn't because they so poor. After she die, he fly to the waterfall in his house. He use lot and lot of balloons [long pause] coming from the chimney. Accidentally, there a boy in front of his house and [long pause] so he has to join the old man too. At first the old man hate the boy but later they love each other. In the journey they discover many secret and fight and come home safe. This cartoon is funny, however, moving. It is about our dream and love and friendship, no matter how old we are.

Second Delivery

I like to watch cartoon in my free time. My favourite cartoon is Up. It a story about an old man. He met his wife when they both little kid. They are married and together until they get old. They always want to go to a great waterfall in South America but because they so poor they

didn't do it. After she die he fly to the waterfall in his house. He use lot of balloons coming from the chimney [long pause]. Accidentally, there a boy in front of his house and [long pause] so he goes with the old man too. At first he hate the boy but later [long pause] in the journey they discover many secret and fight and finally come home safe. This cartoon is funny but also moving. It is about our dream and love and friendship, no matter how old we are.

Third Delivery

My favourite cartoon is Up. It a story about an old man. He met his wife when they are both little kid. They always want to go to a great waterfall in South America but because they so poor they couldn't. After she die he use lot of balloons coming from the chimney [long pause] to fly in his house to the waterfall. There a boy in front of his house and so he goes with the old man too. At first he hate the boy but later in the journey they discover many secret and fight and finally come home safe. This cartoon is funny but also moving.

APPENDIX B

EXAMPLE OF REPEATED MONOLOGUE (TRIMMED SPEECH) IN THE CONSTANT-TIME CONDITION (2/2/2)

First Delivery

Hello. I want talk about my favourite film, The curious case of Benjamin Button. The film is about Benjamin life. He look like eighty year old when he was born but when he grow up he become younger. His life is opposite with human life. His mother [long pause] die and his father left him in [long pause] the nursing home. He grow up there and met Daisy. He look old but Daisy is only sixteen. When he is younger and stronger he travel and work on a ship. He come back when he look younger and handsomer and Daisy and him love each other. They have a daughter but he tell Daisy he must leave because he don't want to be burden for Daisy when he become a child. One day he return like a teenager. Daisy take care of him. Then he became a child and then he die when Daisy is old. I like the film because it very sad and moving and it make me cry. I have watch it many time but I still cry. It about love and sacrifice.

Second Delivery

Hello. I want to talk about my favourite film, The curious case of Benjamin Button. The film is about Benjamin life. His life is strange and it opposite with human life. He look like eighty years old when he was born but when he grow up he become younger. His mother dead when he was born and his father left him in the nursing home. He grow up there and met Daisy. He look old but Daisy is only sixteen. Daisy love to dance. When he is younger and stronger he travel and work on a ship. He come back when he look younger and handsome and Daisy and him love each other. They have a daughter but he must leave he tell Daisy because he don't want to be burden for Daisy when he become a child. One day he return like a teenager. Daisy take care of him when he turn to a child. And then he die when Daisy is old. I like the film because it very sad and moving and it make me cry. I have watch it many time but I still cry. It about love and sacrifice.

Third Delivery

Hello. I want to talk about my favourite film, The curious case of Benjamin Button. The film is about Benjamin life. He look like eighty year old when he was born but then when he grow up he become younger. His life is opposite with human life. His mother dead and his father left him in a nursing home. A black woman take care of him and he live there. He grow up there and met Daisy. He look old but Daisy is only sixteen. Daisy love to dance and is very good dancer. When he is younger and stronger he travel and work on a ship and in a war. He come back when he look younger and handsomer and Daisy and him love each other. They have a daughter but he tell Daisy he must leave because he doesn't want to be burden for Daisy when become child. One day he return and just like a teenager. Daisy take care of him when he became a child. And then he die when Daisy is old. I like the film because it very sad and moving and it make me cry. I have watch it many time but I still cry. It about love and sacrifice. It is very meaningful.