The Effects of Time Pressure on Fluency, Complexity, and Accuracy: A Case Study

Paul Garside

Abstract

Many learners express the desire to be able to speak more fluently. Accordingly, fluency development in L2 learning has attracted growing interest in recent years. One pedagogical activity that has been claimed to promote fluency development is the 4/3/2 activity, whereby learners repeat a monologue under increasing time pressure. In this case study, with one advanced learner, a reduced-time (4/3/2) condition was compared with a consistent-time (3/3/3) condition. Output was analyzed in terms of fluency, syntactic and lexical complexity, and syntactic accuracy. Results showed that repeating a monologue in a reduced-time condition does indeed drive short-term fluency gains, particularly in terms of rate of speech, over and above those produced in a straightforward consistent-time condition. Measures of lexical (but not syntactic) complexity and syntactic accuracy offered limited support for the Trade-off Hypothesis and the claim that the push to speak faster leads to a loss of control over other aspects of speech production.

Keywords: *fluency, complexity, accuracy, 4/3/2 activity*

Introduction

The ability to speak fluently is the goal of many language learners; that is, the desire to express their ideas smoothly and relatively quickly. However, fluency often remains elusive, especially in EFL contexts such as Japan, in which opportunities to use the target language are limited. Within the formal Japanese education system, less emphasis is placed on the development of fluency than receptive skills, partly due to washback effects of the university entrance exams. As a result, many students lack spoken fluency upon entering university, a fact that they are often all too aware of. One pedagogical activity that has been claimed to promote fluency development is the 4/3/2 activity, whereby learners repeat a monologue under increasing time pressure. This activity creates ideal conditions for promoting fluency because, with the need for content generation removed beyond the first iteration, the speaker can focus simply on delivering the content at a faster rate (Nation, 1989). Therefore, the 4/3/2 activity can shed light not only on the nature of fluency development but also on the relationship between fluency and the concomitant aspects of speech production; namely, syntactic accuracy as well as syntactic and lexical complexity.

The current study aims to investigate the effects of this activity by analyzing the output of a single advanced learner of English. Specifically, the results of a reduced-time condition (4/3/2) were compared with those of a consistent-time condition (3/3/3). The output was then analyzed in terms of complexity, accuracy, and fluency.

Literature Review

Complexity, Accuracy, and Fluency

L2 speech production has traditionally been divided into the three main elements of complexity,

accuracy, and fluency (CAF) for the purposes of analysis. Fluency can be further sub-divided into the following three components: rate of speech; breakdown fluency, which refers to the number, duration, and placement of pauses; and repair, which refers to the number of false starts, repetitions, and self-corrections made by a speaker (Tavakoli & Skehan, 2005). Accuracy simply refers to the presence or absence of syntactic errors in an utterance. In terms of complexity, CAF research has traditionally prioritized syntactic complexity, although Skehan (2009) has emphasized the need to include an additional measure of lexical complexity to provide a fuller account of L2 speech production.

In his influential Trade-off Hypothesis, Skehan (2009) has claimed that limitations in the attentional capacity and working memory of L2 learners cause these three aspects of speech production to compete with each other for attentional resources. Specifically, he has maintained that increased fluency can be accompanied by either an increase in accuracy or complexity but not ordinarily by both (Skehan, 2009). In other words, a fundamental tension exists between form on the one hand and fluency on the other. This idea makes intuitive sense, given that increasing time pressure is bound to reduce cognitive processing time and opportunities for online planning. As a language instructor, I believe it also underlines the need for a balanced curriculum, in which sufficient attention is paid to each element of the CAF framework. For example, an overemphasis on fluency can lead to fossilized mistakes if learners are not provided with form-focused corrections and encouraged to reflect on the accuracy of their usage (e.g., Schmidt, 1983). Conversely, a preoccupation with grammatical accuracy can severely hinder fluent output, while also causing learners to avoid some of the more complex features found in well-developed speech, such as relative clauses.

Aspects of the Trade-off Hypothesis, and its claim that complexity, accuracy, and fluency are necessarily in competition have been challenged, chiefly by Robinson (2005) and his Cognition Hypothesis. Whereas Skehan (2009) has stated that, due to limitations in attentional resources and working memory, an increase in task complexity negatively impacts upon all elements of production (CAF), Robinson has claimed that an increase in task complexity can actually facilitate increases in complexity and accuracy, although not fluency. According to Robinson, this is because increasing cognitive demands encourage learners to try to access the more sophisticated linguistic resources required to carry out such tasks successfully. This effort drives learner development as control is gained over the forms and concepts utilized for the task. Crucially, more complex tasks also promote interaction among learners leading to the negotiation of meaning, in which participants pool their resources in an attempt to complete the task. Robinson has also asserted that individual differences play a key role in this dynamic, claiming that learners with higher IQ, aptitude, and working memory actually perform *less* fluently in more complex tasks because of the greater effort they expend to access higher level linguistic resources.

Skehan (2009) responded to this challenge by re-examining some of his earlier studies using a more fine-grained methodology, which included addressing the issue of pause counts within the subcategory of hesitation phenomena. By doing so he was able to make some interesting observations, especially relating to the differences between the timing of native (NS) and non-native speaker (NNS) pauses. He noted that without pre-task planning both NSs and NNSs pause at clause-end boundaries; in fact, NSs actually pause *more* at this point, seemingly regarding it as the natural place for pausing during online planning, or in other words while thinking about what to say next. The main insight, however, is that NNSs pause far more mid-clause, which is what causes speaking turns to lack smoothness. Skehan has suggested that this kind of pause represents a true dysfluency and is caused by NNSs lacking rapid access to the necessary linguistic knowledge. This finding implies that the timing, rather than simply the number or duration of pauses, is important. It also suggests that

fluency is a more nuanced issue than had previously been supposed.

Regarding lexical measures, Skehan (2009) found that more complex vocabulary correlated with more complex syntax for NSs, although it had the opposite effect for NNSs, whose syntax became both less complex and less accurate. He concluded that the extensive lexicons of NSs allow them to access vocabulary relatively effortlessly, which in turn allows parallel processing to continue automatically. L2 learners, on the other hand, require greater effort to access items initially, and then to use them appropriately. This extra cognitive effort has a deleterious effect on both grammatical accuracy and fluency, which clearly demonstrates how trade-offs works in practice.

The 4/3/2 Activity

The 4/3/2 activity was designed as a fluency-building exercise, in which L2 learners are required to repeat a monologue, as close to verbatim as possible, under conditions of increasing time pressure. Repeating the same content within a reduced-time frame (first 4 min, then 3 min, then 2 min) creates ideal conditions for promoting more fluent output because, with the need for content generation removed between the second and third iterations, speakers can focus solely on delivering the content at a faster rate (Nation, 1989).

In psycholinguistic terms, the positive effects of task repetition on fluency are related to the proceduralization, and eventual automatization, of linguistic knowledge (DeKeyser, 2007). Repeating particular lexical and syntactic items has the effect of gradually embedding them in the learner's mind, in association with the function they are used to perform, making them easier to retrieve when called for under similar conditions in future (de Jong & Perfetti, 2011). This process helps to explain the importance of formulaic speech for the development of fluency, as prefabricated phrases that are available for rapid use will naturally allow speakers to assemble units of discourse more quickly than if each phrase has to be reassembled item by item (Wray, 2002). However, it is not clear to what extent, or in what situations, fluency gains (or indeed gains in syntactic accuracy or complexity) made from task repetition transfer to new tasks, which suggests that repetition alone may not be sufficient for acquisition to occur (Ellis, 2009).

Content repetition, and the way in which it promotes fluency development, relates to Levelt's (1989) influential model of speech production. According to this model, speech production depends on the interaction of three separate components: the conceptualizer, the formulator, and the articulator. The content of an utterance is planned in the conceptualizer, therefore if the content is unchanged during a future iteration then less planning time is required, thus facilitating more rapid delivery via the articulator. In theory, greater resources are also made available for the precise wording of the content, which could have additional benefits in terms of complexity and/or accuracy. However, the extent to which these additional benefits are realized during more fluent output represents the nub of the trade-off debate.

The 4/3/2 activity was popularized by Nation (1989), who, in a small-scale exploratory study, found that significant fluency gains were made, especially between the first and third iterations of repeated monologues. The methodology has been much improved upon since, although findings have followed a broadly similar pattern; namely, that increased time pressure promotes more fluent output with, at best, mixed results for accuracy and complexity. For example, Thai and Boers (2016) compared a reduced-time condition with a straightforward repeated task condition. The authors drew on many of the concepts and procedures outlined in Skehan (2009), with complexity, accuracy, and fluency operationalized in similar ways and with the same distinctions made between temporal

and hesitation phenomena, as well as between syntactic and lexical complexity. Their analysis was even more fine-grained, as the output was presented in terms of both raw and trimmed speech (the latter referring to syllables per minute once filled pauses, repeated syllables, and false starts have been excluded), with differential rate of speech measures calculated. A further difference in Thai and Boers' (2016) study was the adoption of the 2,000 most frequent words of English as the threshold for lexical sophistication (Laufer, 2005). However, the most important modification was the inclusion of a *constant-time* comparison group, to be compared with the *shrinking-time* condition. This inclusion allowed the authors to assess the differential effects of reducing the time for each talk as opposed to straightforward task repetition. It was clear from their results that the reduced-time (in this case 3/2/1) condition produced a much higher rate of speech than the consistent-time (2/2/2) condition. Accuracy and complexity were largely unchanged in the former condition, while it increased in the latter condition, although sometimes only slightly. The authors claimed that these results support the Trade-off Hypothesis, although this is questionable as accuracy and complexity did not fall by much, if at all, in the reduced-time condition. I believe it would be more legitimate to say that accuracy and complexity simply did not improve as much as in the consistent-time condition.

Research Questions

The aims of the current study are twofold: first, to identify the relationship between the apparently competing elements of complexity, accuracy, and fluency; and second, to identify the effects of task repetition. Regarding the former, it is essential to measure all three CAF components during several stages of a task to observe any changes that could indicate the shifting allocation of attentional resources. Regarding the latter, if it is indeed the case that repeating key lexical and syntactic patterns under time pressure drives fluency development, as de Jong and Perfetti (2011) have claimed, it needs to be ascertained that time pressure is an essential element of this process. In other words, it needs to be ruled out that repeating those structures without time pressure leads to similar outcomes. It is, therefore, necessary to include both a reduced-time condition and a consistent-time condition, as was the case in Thai and Boers' (2016) study. This is the approach followed here, with an additional focus on hearing the participant's views about experiencing the two conditions. The research questions to be investigated are as follows:

1. Does increased time pressure promote fluency beyond simple task repetition?
2. To what extent do fluency outcomes relate to outcomes in complexity?
3. To what extent do fluency outcomes relate to outcomes in syntactic accuracy?

Participant

The participant was a 43-year-old Japanese female called Yuri (pseudonym). She obtained a TOEIC score of 880 in 2012. As any score above 800 is typically considered in the advanced proficiency range, this clearly identifies her as an advanced-level learner of English. She also has a TOEIC speaking score of 140. Like most recently educated Japanese people, Yuri studied English formally throughout junior high and high school. She then continued for a further 2 years at junior college. Some years ago, she spent a year each living and working in Australia and Canada, which she attributes as the main reason for her high level of English. She has not had as much exposure to English recently, however, and feels both her fluency and confidence have suffered as a result. My own impression was that Yuri's spoken production had remained relatively fluent, although there

were persistent issues with grammatical accuracy.

Procedures

Topic selection is of vital importance when using the 4/3/2 activity. If a topic is either unfamiliar or too demanding (whether linguistically or conceptually), learners are unlikely to make significant gains in fluency. Three topics were selected for the current study, in consultation with the participant, although only two were used. Having an extra topic meant that Yuri could not be sure which topic would be used during the recorded sessions, and also discouraged her from preparing or thinking about the topic in advance.

The *reduced-time* condition was conducted in the first session, with the topic being "Using Tokyo's transport system with a small child." Two question prompts were provided to stimulate ideas and to ensure that she would have sufficient material to talk about. They were: "What is it like using Tokyo's transport system with a small child?" and "How would you like to change it?" Yuri talked about this topic initially for 4 min, then 3 min, then again for 2 min. Ideally, a different interlocutor would be present for each iteration of the talk to alleviate any pressure on the speaker to modify the content in order to maintain the interest of the listener. However, having explained the purpose of the activity, I was satisfied that Yuri was aware of the need to try and repeat her monologue using the same words, and that it was unnecessary to embellish the content for the listener's benefit. I then asked Yuri about her feelings regarding the activity; for example, whether she had enjoyed it, felt any pressure to speak faster, or consciously tried to improve the linguistic aspects her of delivery.

The *consistent-time* condition was conducted 3 days later, with the topic this time being "The way to raise a bilingual child." The two question prompts were: "How can a family raise a bilingual child?" and "What are the advantages and disadvantages of raising a bilingual child?" Once again, Yuri spoke three times but on this occasion each iteration was 3 min in duration to provide a comparison with the reduced-time condition. The purpose of the comparison was to isolate the effects of increasing time pressure from the effects of simple task repetition, although the total amount of speaking time was the same in both conditions (9 min in total). For both topics, Yuri was allowed only 2 min of thinking time before the first iteration of the talk, which I felt was sufficient to access topic-related knowledge.

Finally, a native speaker of English was recorded speaking for 3 min on the bilingualism topic. The purpose of this was simply to provide an approximate baseline figure with which to compare the NNS subject's performance. It has been suggested that some measures of fluency, especially hesitations and false starts, do not vary across L2 proficiency levels, including even very high levels (Baker-Smemoe, Dewey, Bown, & Martinsen, 2014). It would be of interest, therefore, to ascertain whether this even applies to NSs. If so, it would suggest that these features are less an indication of a learner's stage of development than a reflection of individual speech characteristics.

All the monologues (seven in total) were recorded, then processed using PRAAT speech analysis software (Boersma & Weenink, 2013), and finally transcribed. The aim during transcription was to give the reader as authentic an impression as possible, therefore all false starts and repetitions (all marked with a hyphen) have been included, along with non-verbal fillers (in italics) and extended pauses (marked with an ellipsis). These transcriptions appear in full in Appendices A and B.

Measuring Fluency

Taking rate of speech first, this is commonly determined by calculating the mean number of syllables per minute. This raw measure can then be refined by removing all filled pauses, repetitions, and false starts, thus leaving only meaningful syllables and allowing the rate of trimmed speech to be calculated (Lennon, 1990). Only the trimmed measure is considered here as pauses, repetitions, and false starts are counted under breakdown and repair fluency, so to additionally include them in the rate of speech measure would give them undue prominence overall.

Breakdown fluency relates to the number, duration, and placement of pauses (whether silent or containing non-verbal fillers such as *er* and *um*) during speech, while repair fluency is concerned with the number of false starts and repetitions. Regarding breakdowns, it needs to be decided how long a non-verbal gap is before it is counted as a pause; in other words, at what point a pause becomes a sign of dysfluency. This cut-off point has varied among researchers from between 200–400 ms (de Jong & Perfetti, 2011). I have adopted the upper limit of 400 ms in this study as, after listening to samples of learner speech, I believe that a pause of less than 400 ms does not usually indicate a dysfluency and is often nothing more than a pause for breath.

Measuring Complexity

Syntactic complexity is commonly measured in terms of the mean number of clauses per analysis of speech (AS) unit, reflecting the fact that avoiding a string of short, simple sentences is an important feature of L2 development. One AS unit consists of a main clause and any related subordinate or coordinate clauses. For example, "He lived in Tokyo when he was a child" consists of a main clause ("He lived in Tokyo") and a subordinate clause ("when he was a child"). It can be problematic when considering coordinate clauses in spoken production, however, as conjunctions such as *and* or *but* are often followed by extended pauses that suggest the ensuing phrase is not part of the preceding unit, but instead marks a fresh conceptual start. In this study, I have followed the recommendation of Foster, Tonkyn, and Wigglesworth (2000), who have stated that coordinated phrases should be counted as part of the same AS unit unless the initial phrase is marked by rising or falling intonation and followed by a pause of half a second or more. These conditions were satisfied in the following example, "some stations I couldn't find any staff and I was just struggling," which I therefore counted as two clauses within one AS unit.

For lexical complexity, I used the approach (originally proposed by Laufer, 2005) of using the 2,000 most frequently used words in English as a cut-off point, thus offering an indication of the learner's lexical sophistication. Tom Cobb's *Lexical Tutor* website (Cobb, n.d.) was used to track any changes in lexical sophistication between the various iterations of the monologues.

Measuring Syntactic Accuracy

The final aspect of speech production to be considered is accuracy, which is also the simplest to define in relation to target language norms. In CAF research, accuracy is commonly determined by calculating the ratio of error-free clauses to the total number of clauses, and that is the measure utilized here. It should be borne in mind, however, that while accurate production can indicate control over language, it can also result from the avoidance of more challenging structures (Thai & Boers, 2016).

Results and Discussion

Tables 1 and 2 show how the speech performance results changed over the three iterations of both the 4/3/2 and 3/3/3 activities. This includes the three fluency measures (trimmed speech rate, number of pauses, and number of false starts and repetitions), the two complexity measures (clauses per AS unit and lexical sophistication) and the accuracy measure (ratio of error-free clauses). With only one participant, conditions for the use of parametric statistics are not satisfied; consequently, the use of descriptive statistics is appropriate in this case (Turner, 2014). The percentage change measure refers to differences between the first and third iterations only, in order to assess the overall effect of the activity. Table 2 includes the additional results of the NS performance as a footnote.

Table 1
Speech Performance During 4/3/2 Activity

	4 min	3 min	2 min	% change
Trimmed speech rate	118.25	131.00	148.00	25.16
Pauses	20.25	16.33	21.00	3.70
False starts and repetitions	6.50	8.00	7.00	7.14
Clauses per AS unit	2.08	2.35	2.67	28.34
Lexical sophistication	2.02	1.94	1.60	-20.79
Error-free clause ratio	0.68	0.37	0.48	-29.41

Table 2
Speech Performance During 3/3/3 Activity

	1st	2nd	3rd	% change
Trimmed speech rate	102.00	104.67	118.33	16.01
Pauses	21.33	18.67	20.33	-4.69
False starts and repetitions	4.00	6.33	8.67	116.75
Clauses per AS unit	2.94	3.01	3.29	11.90
Lexical sophistication	5.45	4.72	3.87	-28.99
Error-free clause ratio	0.76	0.67	0.54	-28.95

Note. NS results: trimmed speech rate = 190.67; pauses = 13.33; false starts and repetitions = 6.33; clauses per AS unit = 3.33; lexical sophistication = 7.23; error-free clause ratio = 0.97

Fluency

The first research question was: *Does increased time pressure promote fluency beyond simple task repetition?* In terms of pauses (consisting of silences and non-verbal fillers), there was very little change in either version of the activity. Regarding false starts and repetitions, there was little change over the 4/3/2 activity, although a substantial increase was observed during the last iteration of the 3/3/3 activity. This could suggest that the participant had actually become *less* fluent, although it is interesting that the NS participant produced exactly the same number of such dysfluencies (6.33 per 100 words) as in the second iteration of the participant's talk. This finding suggests that, to some extent, false starts and repetitions is a normal aspect of spontaneous speech. It also offers some support to the findings of Baker-Smemoe, Dewey, Bown, and Martinsen (2014), who claimed that these kinds of dysfluencies do not tend to vary across proficiency levels. Judging by the NS performance, the number of false starts and repetitions produced by this advanced learner falls within

a normal range in both activities and was actually surprisingly low during the first iteration of the 3/3/3 activity.

The most noteworthy result is in terms of rate of speech, emphasized by the fact that speech rate (along with mean length of run) has been found to be the best predictor of overall fluency (Kormos, 2006). As mentioned previously, only the trimmed rate (minus false starts and repetitions) was measured in this study, to avoid duplicating the impact of the dysfluencies, and the results for this measure are quite stark. Some increase (16.01%) took place during the 3/3/3 activity, which is perhaps to be expected given that the content was repeated, although a much bigger increase (25.16%) took place over the course of the 4/3/2 activity. These findings are firmly in line with those of Thai and Boers (2016) and suggest that time pressure did indeed push the participant to speak faster and produce more content than straightforward task repetition alone.

A personal impression, based on my experience of teaching and listening to learners of English, was that the 2-min talk from the 4/3/2 activity sounded considerably more fluent than any other talk during either version of the activity. However, some comments from Yuri herself can shed some light on how she experienced the activities. Talking specifically about the 4/3/2 activity, she described how she felt able to speak progressively more smoothly with each iteration:

> The first [time] ... I was thinking what I should say and try to find the words I use. But second time, because I kind of remember what I said, my mind was quite clear. And the third time, I felt it's maybe more comfortable.

She also commented that she tried to reduce the number of filled pauses she made: "First time I felt like *er*, *er*, *er*, like that. But second time I tried not to do that and tried to speak smoothly, and third time more smoothly." According to the data, she was not successful in reducing the number of such pauses, but the fact that she was aware of the need to do so suggests that she was consciously trying to improve her fluency.

Furthermore, when asked to directly compare the two activities, Yuri alluded to the more challenging nature of the reduced-time condition, hinting that pushing herself to speak faster was indeed what drove the fluency gains that were witnessed: "[The 4/3/2] one's more challenging, so it's like quite exciting 'I have to say quickly, I have to say quickly' but [the 3/3/3] one's not so challenging, just trying to remember what I said."

Complexity

The second research question was: *To what extent do fluency outcomes relate to outcomes in complexity?* Regarding syntactic complexity, the mean ratio of clauses per AS unit increased in both the reduced-time and the consistent-time conditions. In the former, a gain of 28.34% (see Table 1) was observed, while in the latter the gain was a more modest 11.90% (see Table 2). This might seem surprising at first glance and runs counter to the findings of Thai and Boers (2016), who found little evidence to suggest that increasing time pressure promotes syntactic complexity. However, closer analysis of the data in this study offers an explanation for the gains in syntactic complexity. In the 4/3/2 activity there are several examples of concepts being repeated but with ideas that were initially separated by distinct pauses, and therefore forming separate AS units, being compressed into one longer unit consisting of multiple clauses. For example, this is one utterance taken from the 2-min version: "Or *erm* they can have the telephone or like a button to call the staff somewhere near there

so *er* for the- the people who have the baby car much easier to call the staff." Contrast that with the corresponding comment from the original 4-min version: "... and- and also if there- there is the ... the *erm* like ph- phone or ... like talking button at the station ... where you need the- you need help from the staff, it's much easier I thought..." Not only does the 2-min version seem more fluent, with fewer repetitions and extended pauses, but the ideas have been tightened into one clear unit of speech. By contrast, in the first iteration the ideas, although related, appear to be in the process of being formulated and are produced separately, with clear gaps between them. This supports the idea that fluency is promoted, at least in part, by the chunking of smaller units of language into larger ones (Kormos, 2006).

A different pattern was observed with regard to lexical sophistication, however. In the reduced-time condition the number of words beyond the 2,000 most frequently used in English decreased from only 2.02 per 100 in the 4-min version down to an even smaller 1.60 in the 2-min version (see Table 1). A similar trend was observed in the consistent-time condition (see Table 2), although the raw number was inflated slightly by the use of the word *bilingual*, which was given in the title and the question prompts, and therefore easily accessible. Interestingly, the NS data was very similar to Yuri's in terms of syntactic complexity, although far higher in terms of lexical sophistication, reflecting the much broader mental lexicon available to L1 speakers (Nation, 2013).

Syntactic Accuracy

The third research question was: *To what extent do fluency outcomes relate to outcomes in syntactic accuracy?* Just one measure was used for accuracy: the ratio of error-free clauses to the total number of clauses produced. In this case there was a clear drop across both conditions. Accuracy fell by 29.41% in the reduced-time condition, although the 3-min iteration was actually the least accurate. It is not clear why such a U-shaped tendency should be evident, although Yuri's comment (above) about finding the reduced-time condition more challenging could explain the large drop in accuracy between the four and 3-min iterations.

The consistent-time condition saw a similar overall drop in accuracy of 28.95, albeit with a gradual decline across the three iterations. A decrease in the former condition could be seen as consistent with the Trade-off Hypothesis, although it is not clear why a drop should also occur during the consistent-time condition, with no pressure to speak faster. These results differ from those of Thai and Boers (2016), who found almost no difference in accuracy across the reduced-time condition and, in fact, a marked improvement across the consistent-time condition. The accuracy data in this study should be interpreted with caution, however, as the errors that Yuri made tended to be very minor, generally relating to articles or prepositions, and not of the sort that would cause comprehension difficulties for listeners. Additionally, it would be interesting to see what kinds of errors are made by lower proficiency speakers during such an activity, and whether they are more disruptive to comprehension.

Conclusion

The data presented here suggest that repeating a monologue in a reduced-time condition does indeed drive fluency gains, at least in terms of rate of speech, over and above those produced in a straightforward consistent-time condition. Similar results have been found among learners of various proficiency levels and L1s (e.g., Boers, 2014; Thai & Boers, 2016). In this case the findings applied to

an advanced-level learner who could already speak English relatively fluently. In pedagogical terms, this finding implies that even advanced learners can benefit from fluency development activities. It would, of course, be interesting to observe whether such gains are held over the longer term, as short-term gains could simply be the result of priming effects (de Jong & Perfetti, 2011). A longitudinal aspect was beyond the scope of the current study, although de Jong and Perfetti (2011) did report that 4/3/2 fluency gains were maintained on a delayed posttest, which they considered evidence of proceduralization.

Lexical complexity and syntactic accuracy declined during the reduced-time condition, which offers some support to the Trade-off Hypothesis and the claim that the push to speak faster leads to a loss of control over other aspects of speech production. However, the fact that this pattern extended to the consistent-time condition was surprising, although Yuri's comment that she saw this activity as less of a challenge could suggest that she was less focused on improving her performance than in the reduced-time condition. The speech rate gains were matched by gains in syntactic complexity, however, and while this too might seem surprising, I believe it can be explained by chunking theories of language, as individual clauses were formed into larger conceptual units (Kormos, 2006). If true, this suggests a positive link could exist between fluency and syntactic complexity, as clauses that initially seem distinct are tightened into one AS unit when the rate of speech increases.

As it involves only one advanced learner of English there are obvious limitations to the current study and it would be beneficial to conduct further, longitudinal research with learners of other proficiency levels. The use of independent raters to give subjective fluency ratings after each iteration would also be of benefit, since the impression made on listeners is an important aspect of speech production and one which cannot be captured by raw data alone. Nevertheless, this study contributes to the growing body of research suggesting that repeated speech activities, incorporating an element of time pressure, are a useful way of developing spoken fluency.

References

Baker-Smemoe, W., Dewey, D. P., Bown, J., & Martinsen, R. A. (2014). Does measuring L2 utterance fluency equal measuring overall L2 proficiency? Evidence from five languages. *Foreign Language Annals*, *47*(4), 707–728. doi:10.1111/flan.12110

Boers, F. (2014). A reappraisal of the 4/3/2 activity. *RELC Journal*, *45*, 221–235. doi:10.1177/0033688 214546964

Boersma, P. & Weenink, D. (2013). *PRAAT: Doing phonetics by computer* (Version 5.3.51) [Computer program]. Retrieved from http://www.praat.org/

Cobb, T. (n.d.). *Range for texts* (Version 3) [Computer program]. Retrieved August 15, 2020, from https://www.lextutor.ca/cgi-bin/range/texts/index.pl.

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

DeKeyser, R. (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge, England: Cambridge University Press.

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*(4), 474–509. doi:10.1093/applin/amp042

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

Kormos, J. (2006). *Speech production and second language acquisition*. New York, NY: Routledge.

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*(3), 387–417.

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

Nation, I. S. P. (1989). Improving speaking fluency. *System*, *17*, 377–384.

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge, UK: Cambridge University Press.

Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics*, *43*, 1–32.

Schmidt, R. (1983). Interaction, acculturation, and the acquisition of communicative competence: A case study of an adult. In M. Wolfson & E. Judd (Eds.), *Sociolinguistics and second language acquisition* (pp. 137–174). Rowley, MA: Newbury House.

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

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: Benjamins.

Thai, C., & Boers, F. (2016). Repeating a monologue under increasing time pressure: Effects on fluency, complexity, and accuracy. *TESOL Quarterly*, *50*(2), 369–393.

Turner, J. L. (2014). *Using statistics in small-scale language education research*. New York, NY: Routledge.

Wray, A. (2002). *Formulaic language and the Lexicon*. Cambridge, UK: Cambridge University Press.

Appendix A
4/3/2 Activity

Four-Minute Iteration

You know the- the *um* the Tokyo transportation system, the train system is really complicated and even though I've been in Japan for years I need to see the maps to travel around Tokyo train ... and ... I found it's more difficult to travel around in Tokyo using the ... trains with a small baby, especially with a baby car ... cos *erm* ... some stations has- have lift but some stations don't so especially when you transfer from one station to the other stations, oh no, one- one- *erm* one line to the other line, *er* sometimes you find the lift all the way to the- the other platform but sometimes you find a lift to just the middle of the changing route, and then no lift after that and- and then sometimes you have to just go up the stairs, and ... I found the difficulties couple of times and ... the new stations has- have more new facilities of course, so it's easier to travel around but sometimes you have to use the old stations as well, and when- when I travelled before ... with those old stations, I had to ask the- the station staff to carry the ... baby car with me, but some stations I couldn't find any- any staff, and I was just struggling and I- I couldn't move from the *er* under the stairs and eventually some ... people offered me to help but until get the- the help you can't actually move, especially if you're child- children sleeping in the baby car it's more difficult. And I- I thought if you *er* if they ... they have the notes say like to change to this line you can use the lift until here or- or you can just say *er* they- they can just say no lift here like that it's- it's much easier I thought and if you can get the information from int- maybe you can get in- get information from the internet but I- I don't know how to check that, so it's- it's more- it's complicated ... and- and also if there- there is the ... the *erm* like ph- phone or ... like talking button at the station ... where you need the- you need help from the staff it's much easier I thought, cos otherwise you just have to wait someone offer you to help. Of course it's difficult to have *erm* make lifts all the stations now cos *er* most of the stations are quite old but so- still at least they can- they could help us with using the notes, I hope.

Three-Minute Iteration

You know the- the Tokyo's the train system is really complicated. Even though I've been in Japan for years I need to look at the map when I go somewhere around in Tokyo, cos from the me- metro to JR or *er* ... private line to metro so complicated. And I found more difficulties to travel around with a small child, especially with a baby car. When I tried to go out travel in Tokyo, *um* some point I couldn't find the lift, and ... *um* ... some- some station has- have lift all the way thro- all the way to the- the other line, but some station don't have the lift cos they're- they're old stations ... and some places have a lift until the middle of the changing route ... and no lift after that, and just stairs or the escalators. I- I had to wait for- I have to ask- I had to ask the staff to help me to carry the baby car go up the s- stairs before, and then if- if I couldn't find a staff I had to wait until someone's help me ... and some ... I- I know some stations can't have lift because they're *er* really old stations, but I hope the- the station have the notes to say you- you have the- you can use the lift from this line to this line, or *er* or you can ... use the lift until here and after that no lift or li- something like that. Maybe you can check that information through internet but I don't know how to use- how to check that, so m- maybe it's difficult for quite a lot of people to find out information. A- Also *erm* if there is *er* the telephone or ... or bell to call the staff *er* near the lift or near the stairs or escalators that's much better because we can ask help- for help, but i- if- if you can't find anyone just- you just have to wait- stand there until y- someone offers you to help,

especially your ch- child, children as- asleep in the baby car, really can't do anything, just standing there, *um*...

Two-Minute Iteration

You know the Tokyo's *er* train system is really complicated. Even though I've been in Japan for years I still need to look at the train map when I travel around in Tokyo. And I found more difficulties to travel around Tokyo with a small child, especially with a bu- *er* baby car, cos *um* some station have lift from one line to the other line, but some station don't have the lift, or just- or have a lift from *er* the line- the platform to the middle of the ch- *um* changing route, and then no lift after that. Sometimes you have to use the stairs or escalators. And with- with a baby car you can't- of- of course you can't just go up *er* s- stairs by yourself. And I have to- I had to ask the train- *er* station staff to help me before. And if you can't find a staff you just have to wait there, and then someone might offer you to help. I hope the- the train station will have the notes saying *er* this station doesn't have a lift from here to the- the other line or like y- there is a lift until here but no lift after that, or like that. Or *erm* they can have the telephone or like a button to call the staff somewhere near there so *er* for the- the people who have the baby car much easier to call the staff. Otherwise you just have to wait standing and th- just stairs until someone offers you. Especially the- if your child is...

Appendix B
3/3/3 Activity

First Iteration

To raise a bilingual child, I think it's- *er, erm* input is very important when they are early stage. *Um*, if they ... hear a lot of language ... they get input and when they have the time becoming speaking they can ... *er* they can use the words they already heard when they were younger. For example, I raise my child as a bilingual, *em* ... I try to speak English to my daughter and ... my daughter speaks Japanese and English now *em, er* what I tried to do was try to ... *er* speak English and also try to make her watch English television. If there are subtitle in English I normally change to English, and then the- the very important thing is *er* to make her think it's natural to hear English, cos if I make her- if- if I make her think she has to learn English she probably refuse it. And I think disadvantage of raising a ... child as a bilingual is when they were- when they are young they probably feel they're different from the others. For example, my daughter w- mi- will- might *er* will feel why she speaks English ... but not the others cos she's gonna go to the Japanese school. So I think *er* that stage probably will be a difficult time but in the future bilingual will be the advantage of them- for them. They can communicate with *er* the people from many countries and there- there are more *er, erm* there are more jobs they can get and ... I think *erm* there- there- they can see more things if they are bilingual *erm*.

Second Iteration

To raise a bilingual child, I think it's *er* important to input the language to the child when they were really ... early stage. For example, I have a daughter and I try to raise her a bilingual in *er* English and Japanese bilingual and then I try to input her a lot of ... English when they were little- *er* w- when she was little and I tr- *um* ... I tried to make her watch English TVs and then I tried to speak English to her at home. So now she's four and she can speak English and Japanese ... and also it's very important thing is *erm* make her think hearing English is natural thing ... and *er* w- whenever there are the program *er* s- English and Japanese I usually change to English subtitles *er* so sub- English and ... *um* ... the- I think the disadvantage of *erm* being a bilingual is when they were young they probably feel they're different from the others. For example, my daughter might feel w- why I'm speaking English but not the others ... but then when they become older it will be- the bilingual will be the advantage because *er* they can ... sp- *er* they can *er* make- they can speak to the- the people from many countries and there are more job- jobs *erm* using English and Japanese *erm* ... *erm serm* but to get to the stage it might be- *er* she mi- the bilingual child might have a difficult to- difficult times ... *erm* ... but to- as I said before it's really important to make them feel it's natural thing to ... speak ... English and Japanese at home or- or whenever so I- I- I'm trying to do that now.

Third Iteration

To raise a bilingual child, I think it's very important to input the language to children when they were young- ... young age and hearing a l- lot of lang- that language make them *er* speak naturally faster as well. For example, I have a daughter and I raise my daughter as a bilingual English and Japanese bilingual and I- *er* when she was little I tried to input a lot of English using TV and als- by also I s- I tried to speak English to her at home and ... I tried to *er* make her watch TV in English if they- the program has English and Japanese I u- usually choose English but the important thing was *er* to make her think the *er* he- hearing English is natural thing so once she he- hears the- hears Japanese she- she say w-

why can't I- can't why watch in Japanese so I ... try to make her watch in English natural way. *Erm* I think *erm* the disva- disadvantage of being a bilingual child is *er* maybe when they- when they were young they probably have the- the feeling they're different from the others and they might think *er* why do I speak English but not the other- *er* but the other ones not speaking English and they- they wanna be the same as the others, so they might have the hard time when they were young stage but in the future when they're getting- when they get older they- they think that's the bi- bi- being bilingual is advantage cos they can *er* speak to many people from the other countries and also they can *er*, *erm* they have more there- there are more jobs to apply for if they are bilingual so *erm* ... the- for parents and for children maybe th- when children young it's- it might have the difficult times but it's *erm* it's I think it's- it's *erm* for the future it's really good to be a bilingual.