tesol QUARTERLY
tesol international association

The Poster Carousel in the ESL Classroom: What Happens to Learners' L2 Fluency During Same and Parallel-Task Repetition?

ANN-MARIE HUNTER
York St John University
York, UK

Abstract

This paper reports on the impact of an English as a Second Language (ESL) speaking activity—*the poster carousel*—on English learners' second language (L2) fluency. Two versions of the poster carousel were developed to observe the effect of talking about (a) the same poster three times (same-task repetition) or (b) three different posters (parallel-task repetition). 46 ESL learners took part, and their performances were audio-recorded, transcribed, and analyzed using PRAAT to detail their L2 utterance fluency. The findings suggest that learners were more fluent during repeated performances in the *same-task repetition* poster carousel group. No changes in fluency were observed for the *parallel-task repetition* poster carousel group. A detailed case study explores these observed fluency increases in the same-task repetition group and generates further hypotheses for empirical exploration. Some observations are made which relate to the selection of fluency measures in L2 fluency research.

doi: 10.1002/tesq.3257

INTRODUCTION

Many second language (L2) learners dream of being able to speak their L2 confidently, and L2 teachers want to help their learners achieve their L2 speaking goals. A substantial body of research

suggests that task repetition—giving learners the opportunity to perform a communicative task once and then do it again—can lead to improved L2 speaking performance, especially in relation to oral fluency, and at least in the short term (see Bygate, 2018).

However, despite this, language teachers do not necessarily incorporate task repetition into their teaching (Tavakoli & Hunter, 2018) possibly because they fear it might impact negatively on learners' motivation (Ahmadian, Mansouri, & Ghominejad, 2017; van de Guchte, Braaksma, Rijlaarsdam, & Bimmel, 2016). Teachers may also worry that any benefits of task repetition are lexis-dependent and therefore specific to a particular task (Suzuki & Hanzawa, 2022). A further methodological concern is that much of the existing task repetition research involves carefully controlled research environments and the use of monologic tasks that may not reflect the realities of L2 classrooms (Foster, 2020).

Given the considerable research interest in task repetition, alongside calls to bridge the gap between SLA research and language teaching practice (e.g., Tavakoli, 2020), it is timely that we pause to reconsider the pedagogical value of task repetition. This study sheds light on task implementation factors that influence the effectiveness of task repetition as a pedagogical tool and asks: what happens to learners' fluency as they engage in repetitive task sequences in the language classroom and—ultimately—is *task repetition actually beneficial?*

Fluency in a Second Language

Levelt (1989) posits that when producing speech, a person first *conceptualizes* the message, then *formulates* the message by drawing on stored linguistic knowledge, before *articulating* the message as overt speech. *Monitoring* also occurs at all stages "to ensure accuracy and appropriateness" (Lambert, Aubrey, & Leeming, 2021, p. 332). Levelt's model has since been adapted to describe the *L2* speech production process (de Bot, 1992; Kormos, 2006; Segalowitz, 2010) and these revised models point to vulnerability points in the speech production process, especially for lower proficiency L2 speakers (Lambert et al., 2021), whereby non-parallel processing impacts on the quality of the overt speech, especially their fluency.

Segalowitz (2010) explains that L2 fluency can be thought of as three interrelated domains. *Cognitive fluency* represents the underlying speed and efficiency of the L2 speech production process. *Perceived fluency* refers to the judgment that a listener makes about a speaker's underlying cognitive fluency, and *utterance fluency* relates to the observable features of overt speech that signify underlying cognitive fluency.

For L2 research, assessment, and teaching purposes, the domain of utterance fluency provides a framework through which researchers and practitioners can attempt to understand the cognitive fluency of L2 learners (Tavakoli & Hunter, 2018).

Speakers can demonstrate (dys)fluency in different ways during speaking, for example, speaking more slowly or pausing more frequently, and SLA researchers have long tended to group these utterance fluency behaviors under three sub-constructs: speed, breakdown, and repair (Skehan, 1998, 2003). Speed is typically measured through articulation rate, while breakdown is measured through frequency and/or length of pauses. Repair is often operationalized as frequency of reformulations, repetitions, and self-corrections. Many L2 fluency studies also adopt *global* measures which straddle two or all three of these sub-constructs (i.e., speech rate).

More recently, L2 fluency research has considered the connection between utterance fluency measures and specific stages of the speech production process. This work attempts to distinguish fluency that relates to L2 knowledge from fluency that relates to natural speaking style and idea shifts (e.g., Peltonen, 2018). For example, in relation to breakdown fluency, many now argue that mid-clause pauses indicate cognitive work at the formulation stage of speech production and is therefore a more L2-specific measure (de Jong, 2016) and pauses which occur around clausal boundaries are considered more suggestive of planning at the conceptualization stage—which happens in both L1 and L2 (Felker, Klockmann, & De Jong, 2019; Lambert et al., 2021; Saito, Ilkan, Magne, Tran, & Suzuki, 2018; Skehan, Foster, & Shum, 2016).

While pausing has been explored in this finer-grained way, less work has been done to pinpoint L2-specific speed and repair measures. That said, some researchers are beginning to investigate the location of sound lengthenings (arguably a measure of speed fluency), sometimes referred to in research as *drawls* (Peltonen, 2018) or *prolongations* (Rohr, 2016; Williams, 2023). Rohr (2016) explains that, as with pausing, L2 speakers are more likely to drawl on syllables that fall within clause boundaries than L1 speakers. However, drawls are not commonly included as a measure of L2 fluency (Williams, 2023) and are especially lacking in L2 classroom studies (see Felker et al., 2019). When it comes to repair, there is some awareness that repetitions may perform a similar function to pauses and drawls in that they represent the speaker's need to buy time while formulating speech. Reformulations and self-corrections are different in that they may represent increased monitoring and *increased* cognitive capacity (Witton-Davies, 2014).

Task Repetition

A pedagogical *task* in language-learning terms can be considered "a goal-oriented activity… that involves meaningful use of language" (Van den Branden, 2016, p. 240). Task *repetition*, therefore, consists of learners performing a task and then performing it again, either immediately or after some time has passed (Bygate & Samuda, 2005). A distinction can be made between same-task repetition, where repeated tasks are identical in terms of structure and content, and parallel-task repetition, where the structure of the task is the same, but the content is different (Lambert et al., 2021).

Same Task Repetition and L2 Fluency

Almost all published studies of same-task repetition report an effect on fluency (Boers, 2014; Bygate, 1996, 2001; Lambert, Kormos, & Minn, 2017; Lambert, Philp, & Nakamura, 2017; Suzuki & Hanzawa, 2022; Wang, 2014; although see also: Fukuta, 2016). While these studies adopt a wide range of fluency measures, this finding appears relatively robust to different contexts, participants, repetition conditions, task type, and proficiency level (Hunter, 2017; Lambert, Kormos, & Minn, 2017; Lambert, Philp, & Nakamura, 2017).

When explaining task repetition effects on utterance fluency, earlier task repetition research tended to foreground the *conceptual planning* argument: the initial performance of the task acts as a dry-run (Bygate, 1996, 2001, 2018) and during a subsequent performance, the conceptual work has been taken care of and the speaker can direct resources to formulation and articulation (Lambert et al., 2021). Recent research has also discussed the impact that *lexical priming* or reuse of constructions (Suzuki, Eguchi, & de Jong, 2022) might have on fluency during task repetition, especially when the repeated performances of the task come soon after each other (Bui, Ahmadian, & Hunter, 2019), because the act of having previously performed a task may mean that a "blue-print" of the utterance is stored in memory (Fukuta, 2016; Lambert, Kormos, & Minn, 2017; Lambert, Philp, & Nakamura, 2017, p. 5).

Seeking to elucidate the underlying processes affected by same-task repetition, Lambert, Kormos, and Minn (2017) and Lambert, Philp, and Nakamura (2017) used a range of specific fluency measures and suggested that stages of the speech production process may be affected during different iterations of the task: conceptualization (as measured by frequency of between-clause pauses) in early iterations

and formulation (as measured by mid-clause pauses) during later iterations. A key question, however, is whether swifter formulation is because of *recently* activated (i.e., primed) language or *regularly* activated (i.e., proceduralized) language. From a pedagogical perspective, a concern is that, if lexical priming is the primary reason for increased fluency during task repetition, any gains may be lexis-dependent and less likely to be transferable to other tasks and contexts.

Parallel-Task Repetition and L2 Fluency

Although studies are far fewer in number, parallel-task repetition which is sometimes referred to as *task type* repetition, has also been found to impact on L2 fluency (De Jong & Perfetti, 2011; Lambert et al., 2021; Pinter, 2005; Sample & Michel, 2014). Parallel-task repetition may offer an attractive option for language teachers because it could be more motivating for learners (Ahmadian et al., 2017), and any fluency gains might be more likely to transfer to novel tasks (Lambert et al., 2021). Increases attested to in Parallel-task repetition studies have typically been attributed to familiarity with the general task procedure (Bygate, 2001) and/or familiarity with abstract rules and *structural* priming (Suzuki et al., 2022).

The Poster Carousel

Many popular ESL/EFL classroom activities such as mingles, fairs or speed dating activities (Bailey & Savage, 1994) are *inherently* repetitive (Gatbonton & Segalowitz, 2005) and are particularly conducive to language learning because the "learner remains involved in genuine communication while engaged in repetition" (Segalowitz, 2010, p. 176). The poster carousel (Lynch & Maclean, 1994) is an example of an inherently repetitive classroom activity, which requires learners to design a poster and then to present their poster to multiple visitors. It has a "particular combination of text input, task structure and learner interaction" (Lynch, 2018, p. 203) which may appeal to teachers working within a task-based or task-supported paradigm. The poster carousel activity also incorporates both pre-task and online planning (Foster & Skehan, 1996) because learners have the chance to discuss the topic as they design their poster and the poster then supports the learners conceptually during performances.

Adopting a case study methodology, Lynch and Maclean (2000, 2001) showed how the quality of the talk of five L2 English learners improved over the six performances of the poster carousel with notable

improvements in "syntactic accuracy" (2001, p. 148), "more precise and less awkward" performance (2001, p. 142), "greater precision in choice of words" (2001, p. 144) and "less vagueness" (2001, p. 151). They also note the "correct fluent use of some language forms after initial difficulty" (2001, p. 144). What is more, the researchers found that the learners in their study responded positively to the poster carousel and were able to perceive benefits for their L2 speaking skills.

The poster carousel described in Lynch and Maclean (2000, 2001) was designed for their teaching context—English for Medical Conferences—but can be easily adapted, making it particularly suited for use in research studies of task factors and L2 fluency within a language classroom.

THE CURRENT STUDY

Overall, then, task repetition leads learners to speak with higher fluency during subsequent task performances, although an important question remains: what underlying processes drive these fluency increases? To date, there has not been a detailed empirical exploration of the poster carousel and its impact on utterance fluency so it is difficult to link holistic performance improvements described in Lynch and Maclean's research with specific stages of the L2 speech production process. Furthermore, given that parallel-task repetition may also foster fluency increases (Lambert et al., 2021) and that learners and teachers may respond more positively to this type of repetition, it is important to explore the pedagogical value of a parallel-task version of the activity. Research questions in the current study were therefore:

1. What is the impact of a same-task repetition version of the poster carousel on adult ESL students' L2 fluency?
2. What is the impact of a parallel-task repetition version of the poster carousel on adult ESL students' L2 fluency?
3. What possible explanations are there for any observed fluency changes during the poster carousel?

METHODOLOGY

Design

Intact classes of EFL learners at a private English language school in Central London were randomly assigned to take part in either a

same-task repetition (STR) poster carousel training session ($n = 24$) or a parallel-task repetition (PTR) carousel training session ($n = 22$) during their normal scheduled English class. In both conditions, participants gave three oral performances (a narrative task)—in the STR group, the three performances related to the same task and in the PTR group they were three different tasks of the same type—and this performance data was subsequently analyzed quantitatively (see dependent variables section below). A mixed between- and within-subjects design was used to look for fluency changes within-subjects across three task performances and between-subjects to compare changes between performances for the two different groups. To answer RQ 3, a full set of annotated performances of a single learner was examined.

Dependent Variables

This study adopts three dependent variables which are traditionally associated with speed, breakdown and repair aspects of utterance fluency, respectively.

Articulation rate. Following De Jong and Perfetti (2011), articulation rate was calculated as the total number of raw syllables produced in the 1-minute sample from each performance that was analyzed, divided by total speaking time (excluding all pauses >250 ms), and multiplied by 60. While drawl measurement would have allowed a more precise indication of speed changes related to formulation, it was not performed on the entire sample because of time constraints but is discussed as part of the case study analysis.

Frequency of mid-clause pauses. Skehan et al. (2016), Huensch and Tracy–Ventura (2017), Lambert, Kormos, and Minn (2017) and Lambert, Philp, and Nakamura (2017) have all suggested that between-clause pauses are related to conceptualization and mid-clause pauses to formulation. Arguably, both filled and silent pauses, and combinations of the two (composite pauses) signal processing delays (Kormos, 2006). *Frequency* of pause, rather than length seems to be most associated with judgments of fluency (Suzuki, Kormos, & Uchihara, 2021). Accordingly, the breakdown fluency measure selected for this study was frequency of mid-clause pauses which was calculated by dividing the total number of mid-clause pauses (filled, silent, and composite >250 ms) by the total sample time and multiplying by 60. A

clause was defined as being either an independent or subordinate clause or sub-clausal unit[1] (Foster et al., 2000).

Frequency of repetitions. In L2 speech analysis, a range of repair behaviors (repetitions; reformulations; self-corrections) are typically bundled together and an overall frequency calculated. However, as Dörnyei and Kormos (1998) explain, while repetitions are likely to represent a need to stall (indicative of greater strain on processing) for time, other repair behavior might be more related to monitoring (indicative of reduced strain on processing). Putting these together might have the effect of canceling one another out. Therefore, repair was calculated as the total number of repetitions *only* divided by total sample time and multiplied by 60.

Participants

46[2] ESL students took part in the study and the age range of participants was 18–42 (M = 20.1 years). First languages spoken were: Arabic (3), Kasakh (1), Portuguese (2), Slovenian (1), Serbian (1), Thai (1), Ukrainian (1), French (8), German (7), Spanish (4), Chinese (2), Japanese (4), Korean (4), Swedish (1), Dutch (1), and Italian (5). The proficiency level of participants was Intermediate (B1, CEFR), and the school allocated new students to a proficiency level based on their performance on an in-house grammar test, a short oral interview, and a writing sample.

Procedure

Through piloting and discussion with class teachers, two new versions of the poster carousel were developed that required learners to create storyboards and talk about extreme life experiences.

[1] *An independent clause* being defined as a clause including a finite verb, *a subordinate clause* consisting of a finite or non-finite verb element with at least one other clause element (Subject, Object, Complement, or Adverbial) and *an independent sub-clausal unit*: "*either* one or more phrases which can be elaborated to a full clause by means of recovery of ellipted elements from the context of the discourse or situation" (Foster et al., 2000: p. 366).

[2] A preliminary power analysis with G-Power recommended a total sample size of 86. This was not a possibility in the current study for reasons of practicality (intact classes) and availability of time for manual analysis. Inferential statistics are presented here with effect sizes and observed power, but caution should be exercised when interpreting the findings.

Same-Task Repetition Carousel Procedure.

1. Teacher divides class into AB pairs and gives each pair an envelope with one of seven different stimulus texts (Appendix S1).
2. Each pair creates a six-frame storyboard on poster paper which illustrates the text they were assigned.
3. Teacher pins posters on the wall around the classroom and each pair stands by their poster (Figure 1).
4. As stay with their posters while the Bs move in a clockwise direction to the next poster. Each A uses their poster to tell story to their B visitor and Bs can ask questions (Figure 2).
5. Steps 5 and 6 are repeated until each A has told the same story 3 times to 3 different visitors (Figure 3).
6. Learners go back to original pairs and swap places. The process is repeated but the Bs will now be telling the story and the As will be visiting other posters.

For the current research, a slightly modified version of the STR procedure described above was used so that learners in the STR condition first listened to a story from another person before repeating *that* story three times rather than presenting their own story three times. This was to because participants would be more familiar with the story for

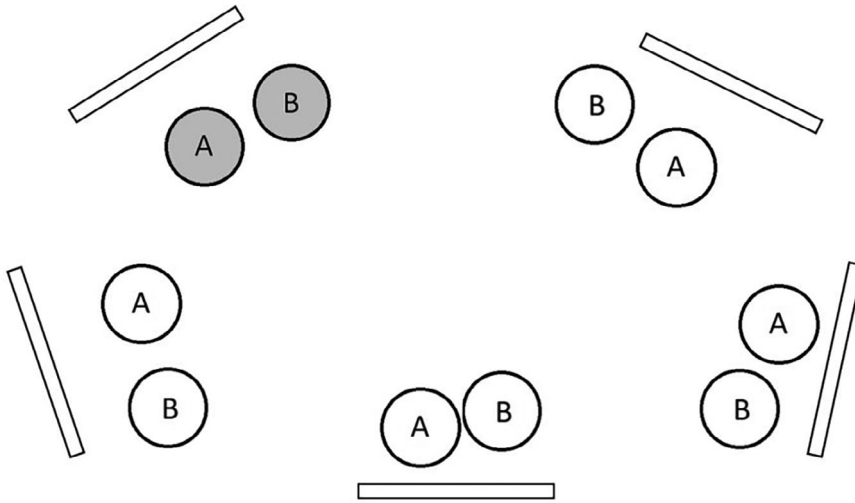

FIGURE 1. Classroom set-up for STR poster carousel with pairs of learners standing in front of their posters. One pair is highlighted for illustrative purposes.

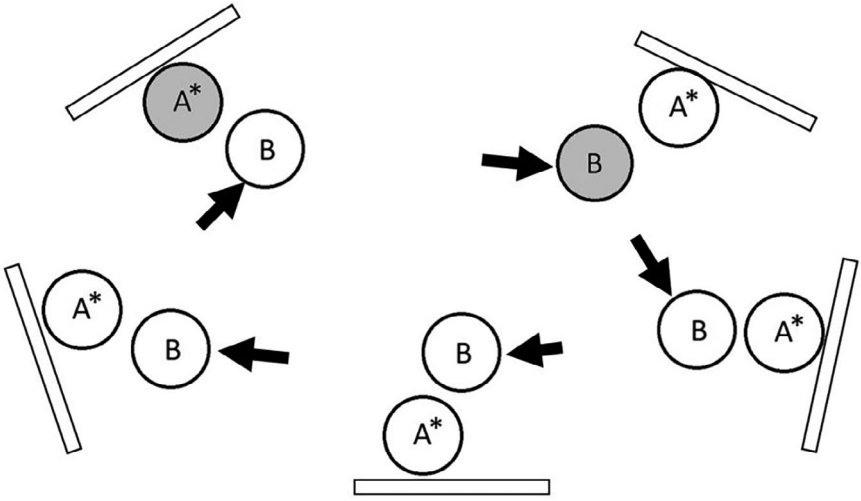

FIGURE 2. Bs moving to next poster. Asterix shows the speaker.

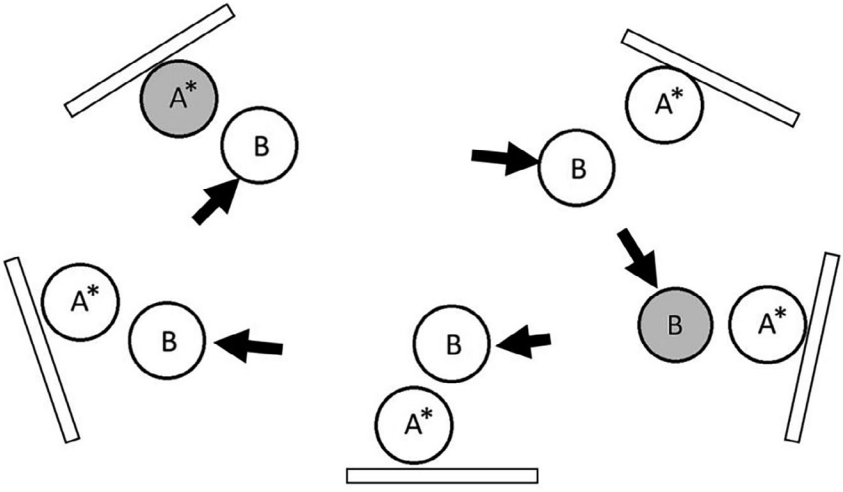

FIGURE 3. Bs moving to the next poster. Asterix indicates the speaker.

which they had designed the storyboard and may have impacted on between-subjects test (described below). The unmodified version is presented above, as this is the version most likely to be used in the language classroom.

Parallel-Task Repetition Carousel Procedure.

1. Teacher divides class into AB pairs and gives each pair an envelope with one of seven different stimulus texts.
2. Each pair creates a six-frame storyboard on poster paper which illustrates the text they were assigned.
3. Teacher pins posters on the wall around the classroom and each pair stands by their poster (Figure 4).
4. As stay with their posters while the Bs move in a counterclockwise direction to the next poster. Each A uses the poster to tell story to their B visitor (Figure 5).
5. Bs and As now SWAP PLACES (Figure 6).
6. As move in clockwise direction to the next poster and Bs present this story to their new visitor (Figure 7).
7. AAs and Bs SWAP PLACES again.
8. Bs move anti-clockwise to the next poster and As present this story to their new visitor.
9. Steps 7–10 are repeated until all As and all Bs have told three different stories.

In the PTR condition, speakers spoke about three different topics, but the procedure followed ensured that the stimulus stories were naturally counterbalanced among the participants to limit the possibility of a task effect.

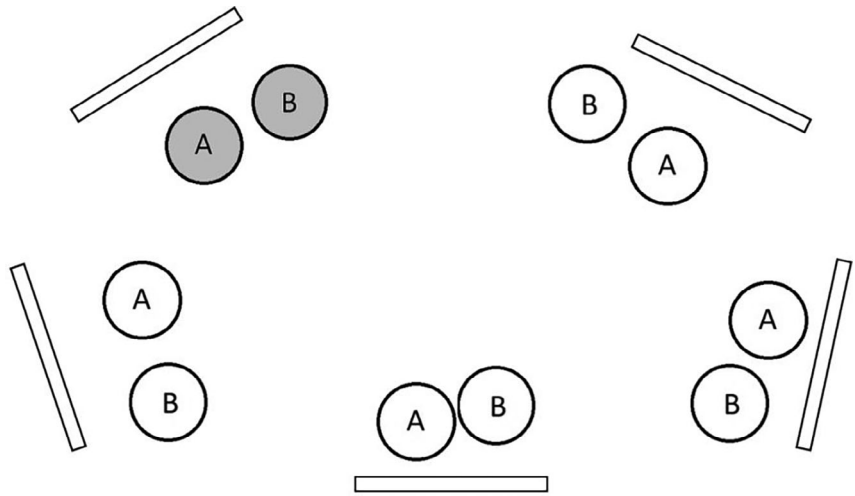

FIGURE 4. Showing classroom set-up at the beginning of the PTR poster carousel. One pair is highlighted for illustrative purposes.

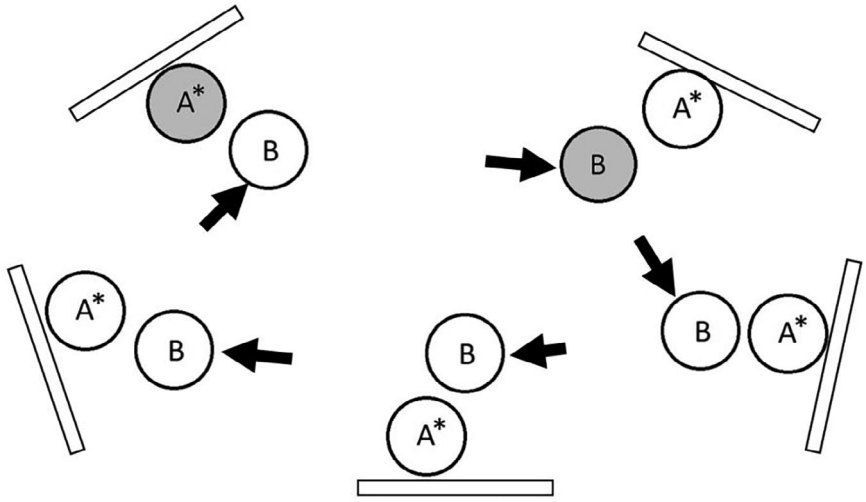

FIGURE 5. Showing Bs moving to the next poster. Asterix indicates the speaker.

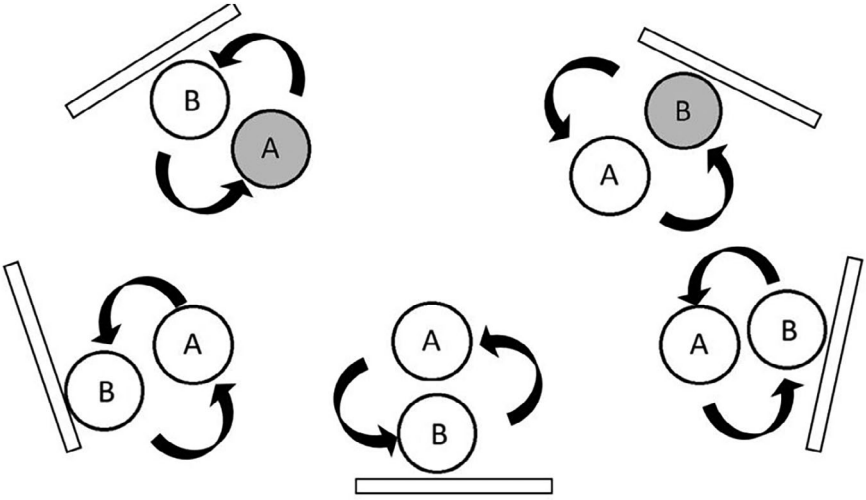

FIGURE 6. Showing As and Bs swapping places.

Materials

The stimulus texts were seven written, first-person descriptions of extreme life experiences (see example text in Appendix S1) from the "Experience" column of the Guardian newspaper (www.theguardian.com/experience): *Cow, Ravine, Baby, Hippo, Cat, Toddler*, and *Railings*.

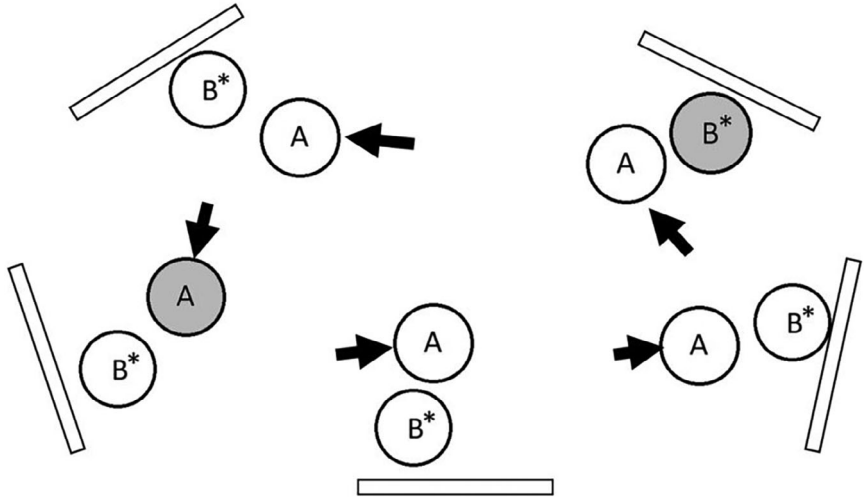

FIGURE 7. Showing As moving in a counterclockwise direction to the next poster. Asterix indicates the speaker.

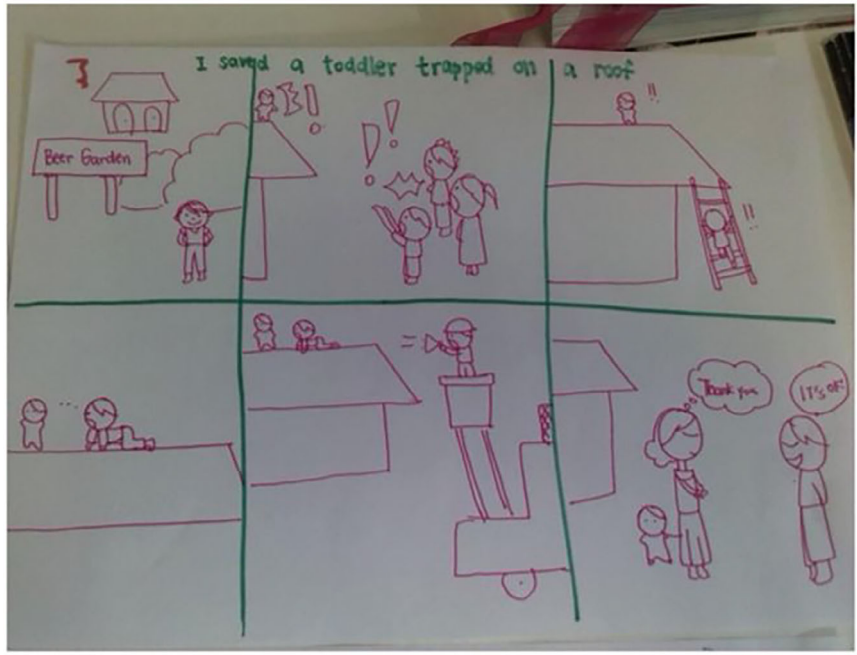

FIGURE 8. Example of storyboard produced by learners.

Careful piloting of the materials revealed that they generated oral accounts that were similar in length and structure. The six-frame, picture-only, storyboards that were created (see Figure 8) supported the students during their retellings. The six-frame storyboard is a highly popular stimulus with a long tradition of use in L2 studies (e.g., Polio & Gass, 1998; Suzuki et al., 2022) and asking learners to create their own storyboards embeds the research within the conventions of a communicative classroom (Lambert, Kormos, & Minn, 2017; Lambert, Philp, & Nakamura, 2017).

Analysis

All three performances for each participant were audio-recorded using Sony ICDBX140 Digital Voice recorders. Participants were given up to 2 minutes to tell their stories, but they did not always use all the allotted time so a 1-minute extract was taken from the beginning of each performance.

Recordings were transcribed both orthographically and using the software PRAAT (Boersma & Weenink, 2013). Automatic fluency analysis (e.g., using PRAAT's *textgrid to silences* feature or De Jong and Wempe (2009) PRAAT syllable nuclei script) was not a viable option in this study because of the levels of background noise as well as the specificity of the fluency measures. Therefore, each extract was annotated *manually* by the author using the standard textgrid feature. Runs of speech, pauses greater than 250 ms,[3] and repair phenomenon were identified through repeated, detailed listening to the recording accompanied by visual inspection of the spectrogram with the screen view zoomed in to, at most, 200 ms (see Figures 9 and 10). Pause position (between- versus mid-clause) and type (silent, filled, composite) were also recorded on the PRAAT textgrids manually by a process of opening the PRAAT file and corresponding orthographic transcription (with clauses marked) simultaneously. Non-verbal phenomena such as laughter, coughing and throat-clearing was discounted from analysis.

A unique PRAAT script was developed which would generate frequencies and durations for the intervals that had been manually created. This output was then used to calculate the specific fluency measures as outlined above.

A 10% sample of the data was re-examined for syllable count (used in the calculation of articulation rate) and pause boundaries by a trained researcher. The second rater coded 22 samples of speech (roughly 10% of the total data) and Spearman's correlations revealed

[3] A commonly used threshold in L2 fluency research.

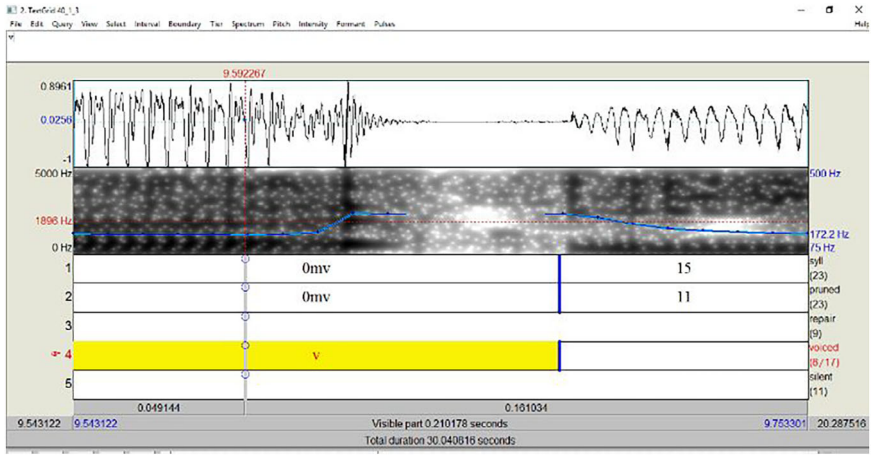

FIGURE 9. Spectrogram with accompanying 5-tier text-grid zoomed to 200 ms.

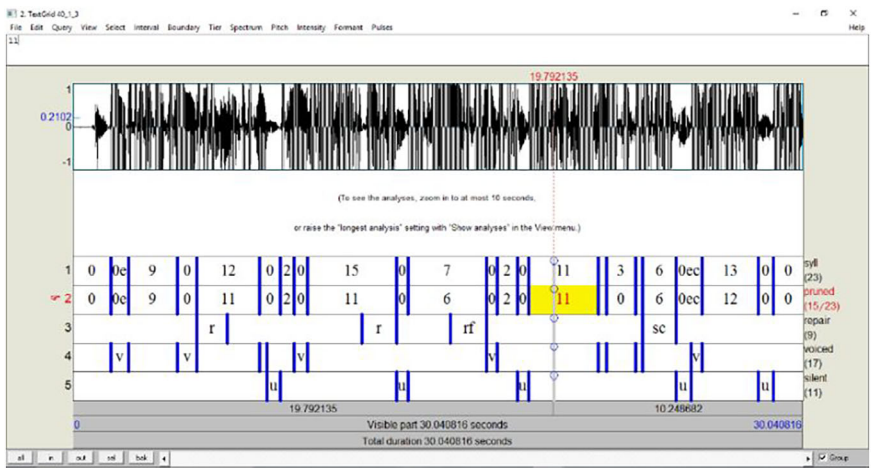

FIGURE 10. 30 s of fully annotated speech.

an exceptionally high reliability score (0.993) between scorers. The time-consuming and specialized nature of manual PRAAT analysis meant that it was not possible to get an inter-rater score for all specific calculations and this needs to be considered when interpreting results.

Statistical Analysis

The data were analyzed using SPSS for Windows. There were no outliers in any cell of the design. In general, dependent variables were

found to be approximately normally distributed as assessed by inspection of Q-Q plots and Levene's test. Some very slight positive skewedness was observed for frequency of mid-clause pauses. However, given that the skewedness was slight and MANOVA is considered robust to this sort of violation (e.g., Bachman, 2004), a decision was taken not to transform the data.

There was homogeneity of covariances, as assessed by Box's test of equality of covariance matrices. Mauchly's test of sphericity indicated that the assumption of sphericity was met for the measures of articulation rate and frequency of repetition. The assumption of sphericity was violated, however, for the measure of frequency of mid-clause pause for the STR group: $\chi^2(2) = 6.428$, $p = .04$. For that reason, the Greenhouse–Geisser adjusted values are reported for this measure.

Effect sizes for the MANOVAs are interpreted following Cohen (1988) who suggests that partial η^2 values of 0.01, 0.09, and 0.25 be considered as small, medium, and large effect sizes, and Cohen's d are interpreted as small ($d = 0.2$), medium ($d = 0.5$), and large ($d = 0.8$). Due to the small sample size, observed power is also reported.

RESULTS

A two-way, repeated measures, multivariate analysis of variance (MANOVA) was conducted to compare within-subjects effects for performance between the two conditions. The within-subjects factor was performance (3 levels) and the between-subjects factor was group (2 levels). Using Wilk's Lambda, the MANOVA revealed that there was a statistically significant interaction for performance and group ($F[6, 172] = 2.942$, $p = .009$; Wilk's $\Lambda = 0.823$, partial $\eta^2 = 0.093$ [a medium effect size], observed power = 0.89). A significant interaction was observed for: articulation rate ($F[2, 46] = 3.997$, $p = .022$, partial $\eta^2 = 0.083$) and frequency of mid-clause pauses ($F[2, 46] = 3.53$, $p = .03$, partial $\eta^2 = 0.074$). There was no significant interaction for frequency of repetition ($F[2, 46] = 1.703$, $p = .188$, partial $\eta^2 = 0.037$).

To answer RQs 1 and 2, which involved establishing the impact on fluency across performances for each version of the poster carousel, follow-up one-way repeated measures MANOVAs were run for each group separately and results are discussed below.

RQ 1. What Is the Impact of a Same-Task Repetition (STR) Version of the Poster Carousel on Adult ESL students' L2 Fluency?

For the STR group, a one-way, repeated measures MANOVA using Wilk's Lambda, revealed a significant within-subject effect for the combined dependent variables ($F[6, 88] = 7.141$, $p < .001$; Wilk's $\Lambda = 0.452$, partial $\eta^2 = 0.327$, observed power = 1) with a very large effect size. Post hoc univariate analysis revealed an effect for performance on articulation rate ($F[2, 24] = 13.036$, $p < .001$, partial $\eta^2 = 0.362$, observed power = 0.996), a trend for Frequency of mid-clause pause ($F[1.543, 24] = 2.999$, $p = .075$, partial $\eta^2 = 0.115$, observed power = 0.482) (although also note the low observed power for this measure), and a significant effect for frequency of repetitions ($F[2, 24] = 7.658$, $p = .001$, partial $\eta^2 = 0.25$, observed power = 0.934) with a large effect size.

Articulation rate. For the same-task repetition group, descriptive statistics reveal that mean articulation rate increased from 214.43 sylls/min during performance 1 to 228.78 sylls/min during performance 2. Mean articulation rate was 227.65 sylls/min during performance 3 (Table 1). Follow-up paired samples *t*-tests with alpha level set at 0.05 revealed significant differences in articulation rate between performance 1 and performance 2 and between performance 1 and performance 3 but not between performance 2 and performance 3 (Table 2) and this is illustrated in Figure 11.

Frequency of mid-clause pause. For the same-task repetition group, a downward pattern is seen across the three performances for frequency of mid-clause pauses (performance 1 = 7.94, performance 2 = 7.02, performance 3 = 6.27) (Figure 12). Paired samples *t*-tests fell short of significance between performances 1 and 2 and between performances 2 and 3. However, the difference between performances 1 and 3 is borderline significant ($p = .057$) (Table 3). The low observed power may be the reason for the weaker significance observed here.

Frequency of repetitions. For the same-task repetition group, frequency of repetitions fell from 4.14 during performance 1 to 2.38 in performance 2. Frequency of repetitions was 2.66 during performance 3. Paired samples *t*-tests revealed significant differences between performance 1 and performance 2 and between performance 1 and performance 3, but not between performance 2 and performance 3. This is illustrated in Table 4 and Figure 13.

TABLE 1
Descriptive Statistics for the Three Dependent Variables of Articulation Rate, Frequency of Mid-Clause Pause and Frequency of Repetitions

									95% confidence intervals	
Dependent variable	Group	Perf.	*n*	Range	Min	Max	Mean	SD	Lower bound	Upper bound
Articulation rate (sylls/min)	STR	1	24	83.18	167.96	251.14	214.43	22.29	202.5903	226.2727
		2	24	104.14	167.80	271.95	228.78	27.65	217.762	239.7997
		3	24	106.29	164.76	271.05	227.65	28.29	216.3814	238.9271
	PTR	1	22	126.28	150.48	276.77	211.63	34.52	199.2576	223.9931
		2	22	100.3	161.97	262.27	211.95	25.80	200.4437	223.4614
		3	22	89.47	167.55	257.01	208.54	26.40	196.7639	220.3121
Frequency mid-clause pauses (per minute)	STR	1	24	13.35	2	15.34	7.94	3.35	6.446939	9.42767
		2	24	12.99	1.2	14.19	7.02	2.81	5.533819	8.51133
		3	24	11.57	3.3	14.87	6.27	2.59	4.852351	7.683107
	PTR	1	22	13.11	3.94	17.05	8.27	3.90	6.710361	9.823633
		2	22	12.58	2.99	15.56	8.66	4.33	7.105733	10.21564
		3	22	14.19	2.91	17.09	9.42	4.18	7.938555	10.89518
Frequency repetitions (per minute)	STR	1	24	13.96	0	13.96	4.14	3.11	2.45466	5.818339
		2	24	8.96	0	8.96	2.38	2.42	0.812286	3.939556
		3	24	10.62	0	10.62	2.66	2.93	1.339218	3.990206
	PTR	1	22	19.83	0	18.83	5.45	4.94	3.695217	7.208465
		2	22	19.77	0.98	20.75	5.15	4.88	3.514527	6.780854
		3	22	12.02	0	12.02	3.83	3.51	2.445063	5.21393

TABLE 2
Pairwise Comparisons STR Group—Articulation Rate (AR)

	Paired differences			95% confidence interval of the difference			
	Mean	Std. deviation	Cohen's *d*	Lower	Upper	*t*	Two-sided *p*
AR1 - AR2	−14.3494	14.62731	−0.981	−20.5259	−8.17278	−4.806	<.001
AR2 - AR3	1.12666	15.73136	0.071619	−5.51611	7.76943	0.351	.729
AR1 - AR3	−13.2227	15.55203	−0.85022	−19.7897	−6.65564	−4.165	<.001

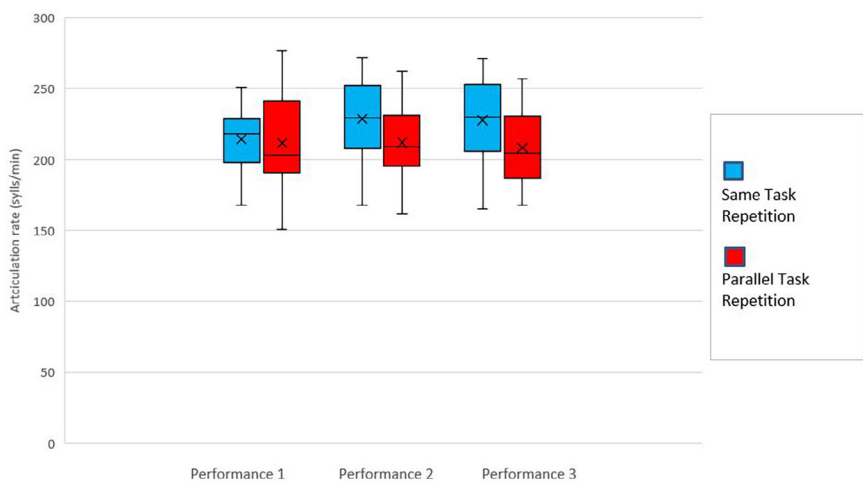

FIGURE 11. Means across performances for articulation rate.

RQ2. What Is the Impact of a Parallel-Task Repetition (PTR) Version of the Poster Carousel on Adult ESL students' L2 Fluency?

For the PTR group, a one-way repeated-measures MANOVA revealed no significant within-subject effect for the combined dependent variables ($F = 1.288$, $p = .272$, partial $\eta^2 = 0.088$, observed power = 0.478). Univariate analysis confirmed no statistically significant effect for performance on articulation rate ($F[2, 22] = 0.229$, $p = .796$, partial $\eta^2 = 0.011$, observed power = 0.084), frequency of mid-clause pause ($F[2, 22] = 1.080$, $p = .349$, partial $\eta^2 = 0.115$, observed power = 0.227), and frequency of repetitions ($F[2, 22] = 2.029$, $p = .144$, partial $\eta^2 = 0.088$, observed power = 0.395). Pairwise

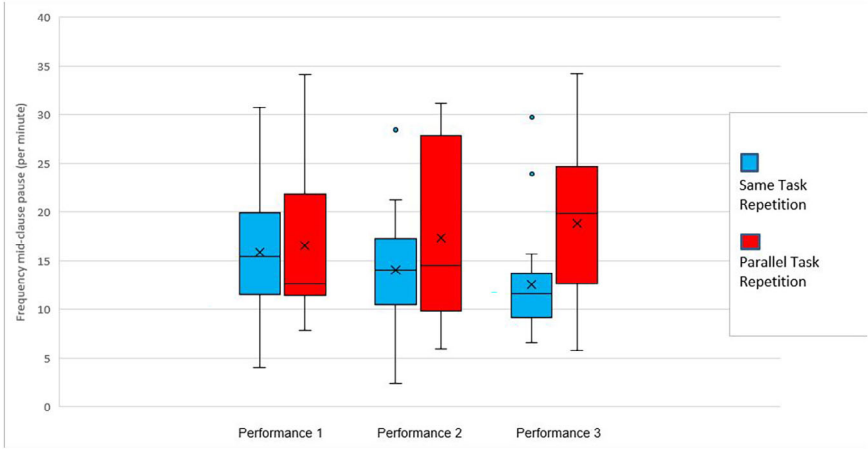

FIGURE 12. Frequency of Mid-Clause Pauses across performances.

TABLE 3
Pairwise Comparisons STR Group—Frequency of Mid-Clause Pause (FMCP)

	Paired differences			95% confidence interval of the difference			
	Mean	Std. deviation	Cohen's *d*	Lower	Upper	*t*	Two-sided *p*
FMCP1 - FMCP2	0.91473	3.29994	0.277196	−0.47871	2.30817	1.358	.188
FMCP2 - FMCP3	0.75485	2.46423	0.306323	−0.28571	1.7954	1.501	.147
FMCP1 - FMCP3	1.66958	4.0745	0.409763	−0.05094	3.39009	2.007	.057

TABLE 4
Pairwise Comparisons STR Group—Frequency of Repetitions (FR)

	Paired differences			95% confidence interval of the difference			
	Mean	Std. deviation	Cohen's *d*	Lower	Upper	*t*	Two-sided *p*
FR1 - FR2	1.76058	2.25756	0.77986	0.80729	2.71386	3.821	<.001
FR2 - FR3	−0.28879	2.75212	−0.10493	−1.45091	0.87333	−0.514	.612
FR1 - FR3	1.47179	2.02338	0.727392	0.61739	2.32619	3.563	.002

comparisons were therefore not carried out for the PTR group but the group means can be observed alongside the STR means in Table 1 and Figures 11–13.

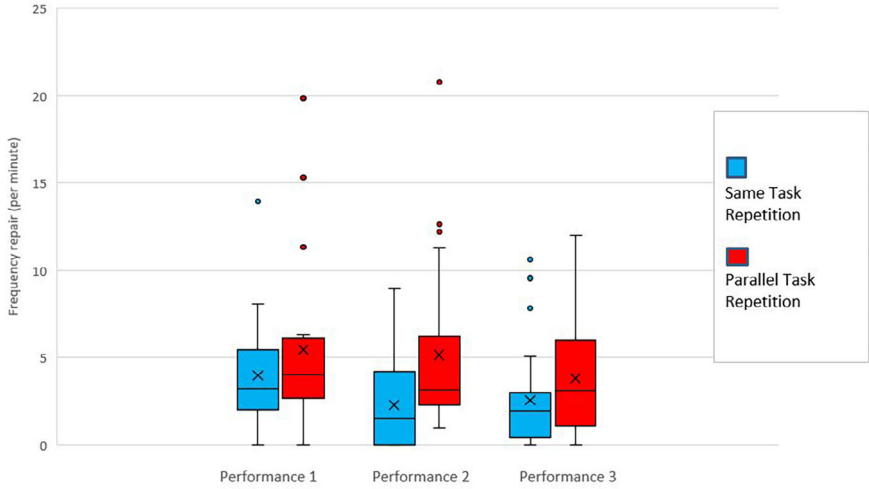

FIGURE 13. Frequency of repetitions across performances.

RQ3. What Possible Explanations Are There for any Observed Fluency Changes during the Poster Carousel?

A nested case study (Kimmy) was selected on the basis of *typical sampling* (Dörnyei, 2007) because Kimmy's utterance fluency profile was largely representative of the STR group overall. A deductive approach was taken to the case study analysis in the sense that the transcriptions (orthographic and PRAAT) were investigated to shine light on the utterance fluency findings (see Appendix S1 for full transcriptions).

Locating Speed Increases. To explore articulation rate changes in this study, Kimmy's performances were analyzed and syllables longer than 250 ms[4] (drawls) were identified through inspection of the PRAAT spectrogram and highlighted on the orthographic transcriptions. All drawls in Kimmy's performances were coded according to their location (mid-clause versus between-clause) and were also measured for length. Details of this analysis can be found in Table 5 alongside the same information relating to Kimmy's articulation rate.

The number of between-clause drawls remain relatively stable across Kimmy's three performances. The number of mid-clause drawls, on the other hand, drops considerably, from 12 during performance 1 to 3 in performance 2. In Kimmy's case, then, some of the increase in

[4] A threshold of 250 ms was chosen because 250 ms is considered meaningful when measuring filled and silent pauses.

TABLE 5
Analysis of Kimmy's Three Performances During Same-Task Repetition Poster Carousel

	Performance 1	Performance 2	Performance 3
Articulation rate (sylls/min)	215.66	228.11	235.66
Number and Av. length of between-clause drawls	10 (0.74 s)	9 (0.71 s)	11 (0.5 s)
Number and Av. length of mid-clause drawls	13 (0.46 s)	3 (0.51 s)	6 (0.49 s)

articulation rate between performance 1 and 2 can likely be explained by this substantial decrease in frequency of mid-clause drawls.

Lexical Priming and Verbatim Repetition. The transcriptions of Kimmy's performances reveal that Kimmy recycles some lexical content from previous performances (Table 6). During narrative detail 1, performance 1, Kimmy lengthens "fo~r" accompanied by a mid-clause pause before the subsequent low-frequency lexical item "charity." During the second performance, there is no hesitation before "charity," suggesting that the word is now primed for use. During narrative detail 5, we see a mid-clause drawl of 0.65 s during the article "a~" which precedes the noun "stick" and another mid-clause drawl during the

TABLE 6
Excerpts at Same Point in Narrative Illustrating Lexical Priming and Verbatim Repetition (Performance 1 versus Performance 2 versus Performance 3)

	Performance 1	Performance 2	Performance 3
Detail 1	the story was about (0.48mf) a woman who worked **for** (0.31ms) a charity (1.28bs)	the story was about a woman who worked for a charity	So the story was about **a** woman (0.88bc) and she was pregnant and worked for a charity (1.73bc) on the countryside
Detail 2	Her friend (0.39mf) who was driving the car	Her friend who was driving the **car** (2.2bc)	N/A
Detail 3	It started to rain very (0.38ms) heavily	It started to rain very heavily	It started to rain very heavily
Detail 4	And dropped (0.27ms) in a ravine (1.27bs)	And dropped (0.26ms) in the ravine	And (0.77bs) drove in a fence
Detail 5	**a** stick of wood (0.31ms) **was** in her chest	a big stick of wood (0.52ms) was in the chest of the pregnant woman (0.94bs)	a big stick of wood **was in the** of **the** preg**nant** (0.65ms) woman's chest (0.88bs)

Note: Pauses are in parenthesis - m = mid-clause; b = between-clause; s = silent; f = filled; c = composite.
Bold indicates drawls.
Mid-clause pauses and drawls are underlined.

verb "wa~s" which precedes the prepositional phrase "in her chest". During the second performance, there are no drawls and the phrases "a big stick of wood" and "was in the chest of the pregnant woman" are uttered without significant hesitation. This suggests that lexical priming may be involved in fluency increases in Kimmy's case.

In detail 3, we see how a complete phrase is re-used in each of the subsequent performances. However, it is also noticeable that Kimmy does not engage in very large stretches of verbatim repetition and she seems equally likely to play with the order of elements in her narrative (detail 1, performance 3), abandon some elements (detail 2, performance 3) and introduce others (detail 4, performance 3). As she does this, we see hesitation around clause boundaries, which is in line with other studies that have considered the impact of (re)conceptualisation on L2 fluency (e.g., Felker et al., 2019) but a reduction in mid-clause hesitation. This suggests that whether language is re-used or not, the speaker is still finding it easier to formulate her utterances in subsequent performances.

Hesitation Clusters and Problematic Language. The analysis of Kimmy's performances highlights particular moments of high dysfluency where drawls co-occur with other hesitation phenomena like filled and silent pauses in hesitation *clusters* (Figure 14).

The lengthiest clusters tend to occur around lower frequency or otherwise problematic language. Table 7 highlights how Kimmy seems to struggle with her explanation of the car leaving the road. This is evidenced by the mid-clause hesitation cluster around verb + preposition "went away" (performance 1) and "get away" (performance 2)

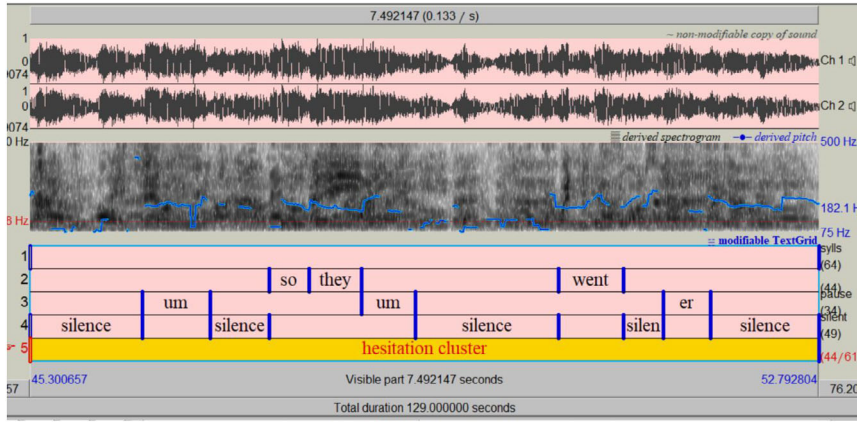

FIGURE 14. Hesitation cluster during Kimmy's first performance.

TABLE 7
Example from Same Point in Narrative Across the Three Performances

Performance 1	Performance 2	Performance 3
it started to rain very (0.38ms) heavily **and (0.82)** it **was (0.52)** (0.4bs) the weather was so bad that her **friend (0.81)** (0.28mf) could not see the road anymore (0.86bs) **and** (0.7) could not control **the (0.32)** steering (1.29ms) wheel yeah (2.32bc) **so (0.38) they (0.5)** (1.85mc) **went (0.62)** (1.87mc) away from **the (0.27)** road or the because she could not see the road (0.43bf) and dropped (0.27ms) in a ravine (1.27bs)	**and (0.81)** (0.55bf) her friend who was driving the **car (0.57)** (2.2bc) had problems to see the road because the weather was very bad so it started to rain very heavily **and (0.78)** it wasn't (0.35ms) yeah possible to see the road so **they (0.41)** (3.3mc) get away from the road or dropped (0.26ms) in the ravine	but then the (0.42) weather was very bad it started to rain very heavily **so (0.41)** (0.38bs) her **friend (0.7)** (0.33mf) could not see the road anymore **and (0.45)** (0.77bs) drove in a fence

Note: Pauses are in parenthesis - m = mid-clause; b = between-clause; s = silent; f = filled; c = composite.
Bold indicates a drawl.
Between-clause pauses and drawls are underlined.

suggesting that she is unsure of the target construction (presumably, "went off" or "veered off"). What is interesting is that, during the third performance, this detail is abandoned altogether. Other constructions which cause mid-clause hesitation are abandoned after the first performance, such as the detail regarding the steering wheel (Table 7).

On the other hand, Kimmy's third performance includes a hesitation cluster which precedes the use of the genitive case, the first time that Kimmy has attempted this structure across the performances (Table 6, detail 5, performance 3). It is possible that the mid-clause drawling seen here relates to the processing and production of that more complex structure, suggesting that not all problematic language is abandoned in repeated performances.

Summary of Case Study. Speed increases in Kimmy's case may be related to a reduction in drawls in the mid-clause position, suggestive of swifter *formulation* during subsequent iterations. Some of this swifter formulation may be related to recently activated (primed) language (e.g., the words "charity," "heavily," "stick"). However, the case study also revealed Kimmy might be consciously selecting language which is more familiar to her (proceduralized language). Kimmy also seems to experiment with more complex language during the final performance.

GENERAL DISCUSSION

Same-Task Repetition Poster Carousel and Fluency

This study found that articulation rate increased between same-task repetition performance 1 and performance 2. Articulation rate is a crude measure, potentially linked to both conceptualisation and formulation. The articulation rate gains for Kimmy corresponded with a considerable reduction in drawls in the mid-clause position between performances 1 and 2, suggesting an association with the *formulation* stage of the speech production process. Interestingly, the speed of articulation did not increase further between performances 2 and 3. There is some evidence to suggest that this could be associated with increased experimentation with language during performance 3 leading to dysfluency.

The only other study which has looked at same-task repetition and its effect on articulation rate is De Jong and Perfetti (2011). They found that articulation rate increased significantly during same-task repetition, but the largest gains were between performances 2 and 3 which is in contrast with the findings of the current study. An explanation could be that it was only the final 2-minute performance which constituted time pressure for the learners in their study. An alternative explanation is that the tasks placed different demands on participants; the fact that the task in the current study was a narrative task as opposed to an opinion task, and perhaps the specific combination of repetition and planning inherent in the poster carousel made earlier increases in articulation rate more likely.

Previous studies have argued that speed increases during task repetition could be linked primarily to speakers reusing constructions (Suzuki et al., 2022; Thai & Boers, 2016) and there is evidence to suggest this is also the case here. However, the case study suggested that the speaker was also reconceptualising the message, abandoning certain elements, playing with order and even experimenting with more complex language. In the poster carousel, participants may be benefitting from opportunities to reuse some constructions but may also reinvest increases in cognitive fluency to monitor their speech and use the information to make strategic changes during subsequent performances.

Frequency of mid-clause pauses reduced gradually over the three performances resulting in a borderline significant decrease between performances 1 and 3. This result mirrors that of Lambert, Kormos, and Minn (2017) and Lambert, Philp, and Nakamura (2017), who similarly found a step-wise reduction in frequency of mid-clause pause.

The current study adds to our understanding of pausing during task repetition with the indication that mid-clause pauses tend to appear around problematic language such as in the middle of phrasal verbs.

In the current study, there was a reduction in the frequency of repetitions between the first and second performances, following a similar trajectory to articulation rate. This is in contrast with other studies which have found no effect for same-task repetition on repair (e.g. Wang, 2014). Lambert, Kormos, and Minn (2017) and Lambert, Philp, and Nakamura (2017) also found no effect for task repetition on repair (frequency of overt self-corrections) over the first three performances of the tasks, although they did find a reduction in the need to repair on subsequent task performances. A possible explanation for the differences in findings here is that this study used a repair measure that targeted stalling as opposed to monitoring behavior.

Parallel-Task Carousel and Fluency

In the current study, and in contrast to a number of other studies (De Jong & Perfetti, 2011; Lambert et al., 2021; Pinter, 2005; Sample & Michel, 2014), there was no significant change in participants' fluency during the parallel-task version of the poster carousel. This might suggest that it is primarily the content repetition, or lexical priming, during same-task repetition that drives the increases in fluency that were observed in this study. Fluency gains for parallel-task repetition in De Jong and Perfetti (2011) study may have been because the participants in their study were performing their tasks under increasing time pressure and participants had less time to fill with speech in later performances.[5] In Sample and Michel (2014) investigation of parallel-task repetition and oral performance, gains in fluency could be attributed to content repetition because the repeated versions of the task were actually *identical* in terms of the lexis they required of students. Lambert et al. (2021) found that parallel-task repetition was associated with almost identical speed fluency gains as the same-task repetition condition. However, they used a global measure of fluency—speech rate—while the current study opted for a pure speed measure (articulation rate). It is possible, then, that speech rate captured changes (e.g. between-clause pauses) not detected by the measures used in the

[5] In the 4-3-2 studies reported in De Jong & Perfetti (2011), the researchers analyzed the whole performance each time. 4 minutes is a long time to fill when answering a single question and so it is likely that those initial performances include lots of filler. It is understandable that a 2 minute sample of speech for a single task will be more fluent than a 4 minute sample of speech.

current study and which may be associated with task familiarity/structural priming.

An alternative explanation is that, because the parallel-task carousel involved passing stories from one participant to the next, details of the narrative may have been omitted leading to conceptual difficulties during the retelling. In addition, the interlocutors in the parallel-task group tended to ask more questions. This is likely because they knew they would have to tell the same story and wanted to be sure they understood the details. It meant that speakers in the parallel-task group were less likely to launch into a prepared narrative and had to (re)conceptualize the message, potentially creating further dysfluency (Felker et al., 2019).

Implications

The study has shown that the poster carousel can be a useful tool to explore task repetition effects in the classroom context and echoes findings of other research into the impact of same-task repetition on L2 fluency, thereby providing support for the use of this inherently repetitive task sequence in the language classroom. Hopefully teachers are inspired to experiment with the carousel technique and explore task factors, learner outcomes, and attitudes further.

The current study has demonstrated that articulation rate alone can be rather blunt, but when taken alongside an analysis of *drawl length* and *location*, provides a more nuanced picture of underlying cognitive fluency. Similarly, while many studies combine a range of repair behaviors for fluency analysis purposes, the findings of the current study add support to the hypothesis that repetitions may perform a very different function to other repair measures. This may provide an explanation as to why a growing number of studies have found that repair measures do not correlate with subjective ratings of fluency (Bosker, Pinget, Quené, Sanders, & De Jong, 2013, Tavakoli et al., 2020; Suzuki et al., 2021) and that studies of task conditions seem to have contradictory findings when it comes to effects on repair. Indeed, there is support here for an approach to L2 utterance fluency analysis which groups dependent variables in terms of the conceptualisation, formulation and monitoring stages of speech production (Lambert et al., 2021) as opposed to speed, breakdown and repair.

Finally, numerous studies have sought to capture longer term or *transfer* fluency benefits of task repetition. The case study presented here has shown that increased fluency *during* task repetition may not be expected to translate into longer term increased *fluency* but may

have been invested in other ways, giving rise to increased complexity or accuracy in the longer term, for example.

Limitations and Directions for Research

Firstly, and similarly to Suzuki et al. (2022), the sample size used in this study ($n = 46$; $n = 22$ for one-way analyses) was smaller than recommended in the a priori power analysis performed using G-Power. Care has been taken to present observed power and effect sizes, but caution is advisable when interpreting inferential statistics. Manual fluency analysis of classroom data is time-consuming, but replications could be carried out which make use of existing PRAAT functions and specially created scripts to measure fluency automatically (De Jong & Wempe, 2009) on a larger scale, perhaps in conjunction with multiple nested case studies which provide more targeted fluency data.

Crucially, further empirical research is needed to test hypotheses generated here in relation to location of speed increases during task repetition, how formulation relates to both lexical priming and strategic message changes as well as whether the specific task demands of the poster carousel are responsible for some of the findings. Finally, it would be valuable to explore learners' perspectives on their fluency increases during the poster carousel activity. This could be achieved through stimulated recall and post-intervention interviews, for example (Kahng, 2014).

Finally, carrying out research in an authentic classroom and with genuinely communicative tasks means that the conditions cannot be as controlled as in a laboratory environment. The tasks used in this study were planned very carefully for internal validity however task factors and group dynamics may have played a role in how these interventions impacted on fluency. Future research might seek to disentangle some of the threads discussed here.

CONCLUSION

This paper has reported on a classroom-based study of the poster carousel and L2 fluency. The study suggested that the immediate same-task repetition inherent in the poster carousel allows learners to perform tasks with increased oral fluency. For this particular classroom activity, it seems that the fluency changes are more likely related to reductions in the time needed to formulate content and this may be due in part to lexical priming and also to the learners selecting language which is more familiar. The parallel-task repetition version of

the poster carousel did not seem to lead to higher L2 fluency in learners, although the specific nature of the parallel-task repetition carousel could have been an intervening factor. It is hoped this study underlines the value of classroom-based enquiry, and the use of authentic tasks in L2 fluency research.

ACKNOWLEDGMENTS

I would like to thank the three anonymous reviewers for their comments on earlier versions of this article.

THE AUTHOR

Ann-Marie Hunter is a Lecturer in TESOL and Linguistics at York St John University. Her primary research focus is L2 speaking and language teaching. She volunteers as an ESOL teacher and volunteer coordinator and is interested in making second language learning inclusive and accessible.

REFERENCES

Ahmadian, M. J., Mansouri, S. A., & Ghominejad, S. (2017). Language learners' and teachers' perceptions of task repetition. *ELT Journal*, *71*(4), 467–477.

Bachman, L. F. (2004). *Statistical analyses for language assessment book*. Cambridge: Cambridge University Press.

Bailey, K. M., & Savage, L. (1994). *New ways in teaching speaking. New ways in TESOL series: Innovative classroom techniques*. Alexandria, VA: TESOL.

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

Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer (Version 5.3.04). Retrieved from http://www.praat.org/

Bosker, H. R., Pinget, A. F., Quené, H., Sanders, T., & De Jong, N. H. (2013). What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing*, *30*(2), 159–175.

Bui, G., Ahmadian, M., & Hunter, A. (2019). Spacing effects on repeated L2 task performance. *System*, *81*, 1–13.

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. 134–146). London: Heinemann.

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

Bygate, M. (Ed.). (2018). *Learning language through task repetition* (Vol. *11*). Amsterdam: John Benjamins Publishing Company.

Bygate, M., & Samuda, V. (2005). Integrative planning through the use of task repetition. In *Planning and task performance in a second language. Language learning and language teaching* (Vol. *11*, pp. 37–74). Amsterdam: John Benjamins. ISBN 9027219613.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
Dörnyei, Z., & Kormos, J. (1998). Problem-solving mechanisms in L2 communication. *Studies in Second Language Acquisition*, *20*(3), 349–385.
Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.
De Bot, K. (1992). Applied linguistics. *Applied Linguistics*, *13*(1), 1–24.
De Jong, N., & Perfetti, C. A. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language Learning*, *61*(2), 533–568.
De Jong, N. H. (2016). Predicting pauses in L1 and L2 speech IRAL (aop).
De Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*, *41*(2), 385–390.
Felker, E. R., Klockmann, H. E., & De Jong, N. H. (2019). How conceptualizing influences fluency in first and second language speech production. *Applied PsychoLinguistics*, *40*(1), 111–136.
Foster, P. (2020). Oral fluency in a second language: A research agenda for the next ten years. *Language Teaching*, *53*(4), 446–461.
Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, *18*(3), 299–323.
Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied linguistics*, *21*(3), 354–375.
Fukuta, J. (2016). Effects of task repetition on learners' attention orientation in L2 oral production. *Language Teaching Research*, *20*(3), 321–340.
Gatbonton, E., & Segalowitz, N. (2005). Rethinking communicative language teaching: A focus on access to fluency. *The Canadian Modern Language Review*, *61*(3), 325–353.
van de Guchte, M., Braaksma, M., Rijlaarsdam, G., & Bimmel, P. (2016). Focus on form through task repetition in TBLT. *Language Teaching Research*, *20*(3), 300–320.
Huensch, A., & Tracy–Ventura, N. (2017). L2 utterance fluency development before, during, and after residence abroad: A multidimensional investigation. *The Modern Language Journal*, *101*(2), 275–293.
Hunter, A. M. (2017). *Fluency development in the ESL classroom: The impact of immediate task repetition and procedural repetition on learners' oral fluency* (Doctoral dissertation, St Mary's University, Twickenham).
Kahng, J. (2014). Exploring utterance and cognitive fluency of L1 and L2 English speakers: Temporal measures and stimulated recall. *Language Learning*, *64*(4), 809–854.
Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.
Lambert, C., Aubrey, S., & Leeming, P. (2021). Task preparation and second language speech production. *TESOL Quarterly*, *55*(2), 331–365.
Lambert, C., Kormos, J., & Minn, D. (2017). Task repetition and second language speech processing. *Studies in Second Language Acquisition*, *39*(1), 167–196.
Lambert, C., Philp, J., & Nakamura, S. (2017). Learner-generated content and engagement in second language task performance. *Language Teaching Research*, *21*(6), 665–680.
Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Lynch, T. (2018). Perform, reflect, recycle: Enhancing task repetition in second language speaking classes. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 193–222). Amsterdam: John Benjamins.
Lynch, T., & Maclean, J. (1994). The poster carousel. In K. M. Bailey & L. Savage (Eds.), *New ways in teaching speaking. New ways in TESOL series: Innovative classroom techniques* (p. 306). Alexandria, VA: TESOL press.
Lynch, T., & Maclean, J. (2000). Exploring the benefits of task repetition and recycling for classroom language learning. *Language Teaching Research*, *4*(3), 221–250.
Lynch, T., & Maclean, J. (2001). Effects of immediate task repetition on learners' performance. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks, second language learning, teaching and testing* (pp. 99–118). Harlow: Longman.
Peltonen, P. (2018). Exploring connections between first and second language fluency: A mixed methods approach. *The Modern Language Journal*, *102*(4), 676–692.
Pinter, A. (2005). Task repetition with 10-year-old children. In Edwards, C., & Willis, J. R. (Eds.), *Teachers exploring tasks in English language teaching* (pp. 113–126). Basingstoke: Palgrave Macmillan UK.
Polio, C., & Gass, S. M. (1998). The role of interaction in native speaker comprehension of nonnative speaker speech. *The Modern Language Journal*, *82*(3), 308–319.
Rohr, J. L. (2016). *Acoustic and Perceptual Correlates of L2 Fluency: The Role of Prolongations* (Doctoral dissertation).
Saito, K., Ilkan, M., Magne, V., Tran, M. N., & Suzuki, S. (2018). Acoustic characteristics and learner profiles of low-, mid-and high-level second language fluency. *Applied PsychoLinguistics*, *39*(3), 593–617.
Sample, E., & Michel, M. (2014). An exploratory study into trade-off effects of complexity, accuracy, and fluency on young learners' oral task repetition. *TESL Canada Journal*, *31*, 23–46.
Segalowitz, N. (2010). *Cognitive bases of second language fluency*. England, UK: Routledge.
Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
Skehan, P. (2003). Task-based instruction. *Language Teaching*, *36*(1), 1–14.
Skehan, P., Foster, P., & Shum, S. (2016). Ladders and snakes in second language fluency. *International Review of Applied Linguistics in Language Teaching*, *54*(2), 97–111.
Suzuki, S., Kormos, J., & Uchihara, T. (2021). The relationship between utterance and perceived fluency: A meta-analysis of correlational studies. *The Modern Language Journal*, *105*(2), 435–463.
Suzuki, Y., Eguchi, M., & de Jong, N. (2022). Does the reuse of constructions promote fluency development in task repetition? A usage-based perspective. *TESOL Quarterly*, *56*(4), 1290–1319.
Suzuki, Y., & Hanzawa, K. (2022). Massed task repetition is a double-edged sword for fluency development: An EFL classroom study. *Studies in Second Language Acquisition*, *44*(2), 536–561.
Tavakoli, P. (2020). Making fluency research accessible to second language teachers: The impact of a training intervention. *Language Teaching Research*, *27*, 1362168820951213.

Tavakoli, P., & Hunter, A. M. (2018). Is fluency being 'neglected'in the classroom? Teacher understanding of fluency and related classroom practices. *Language Teaching Research*, *22*(3), 330–349.
Tavakoli, P., Nakatsuhara, F., & Hunter, A. M. (2020). Aspects of fluency across assessed levels of speaking proficiency. *The Modern Language Journal*, *104*(1), 169–191.
Thai, C., & Boers, F. (2016). Repeating a monologue under increasing time pressure: Effects on fluency, complexity, and accuracy. *TESOL Quarterly*, *50*(2), 369–393.
Van den Branden, K. (2016). Task-based language teaching. In G. Hall (Ed.), *The Routledge handbook of English language teaching* (pp. 238–251). England, UK: Routledge.
Wang, Z. (2014). On-line time pressure manipulations: L2 speaking performance under five types of planning and repetition conditions. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 27–62). Amsterdam: John Benjamins.
Williams, S. (2023). *Disfluency and proficiency in second language speech production*. London: Springer Nature.
Witton-Davies, G. (2014). The study of fluency and its development in monologue and dialogue. Unpublished doctoral dissertation, Lancaster University, Lancaster, UK.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Example of stimulus materials given to learners.

Appendix S1. Case_study_transcriptions_annotated.