| | |
|---|---|
| AUTHOR | Pauliina Peltonen and Pekka Lintunen |
| TITLE | Teaching L2 speech fluency with learner-corpus-based awareness-raising activities |
| YEAR | 2024 |
| DOI | https://doi.org/10.1075/ijlcr.00040.pel |
| VERSION | Author accepted version |
| CITATION | Peltonen, P. & Lintunen, P. (2024). Teaching L2 speech fluency with learner-corpus-based awareness-raising activities: Insights from a short-term intervention study. International Journal of Learner Corpus Research 10 (1), 39–66. Special issue "Learner Corpus Research for Pedagogical Purposes" (edited by S. Götz & S. Granger). DOI: https://doi.org/10.1075/ijlcr.00040.pel |

**Teaching L2 speech fluency with learner-corpus-based awareness-raising activities: Insights from a short-term intervention study**

Pauliina Peltonen and Pekka Lintunen

University of Turku

Second language (L2) speech fluency has been found to develop especially in study-abroad contexts, while few studies have examined fluency development in formal instruction. Furthermore, the instructional approaches have rarely targeted fluency specifically. The present study examines the effects of a teaching experiment utilizing learner-corpus-based awareness-raising activities on L2 fluency development in formal instruction. Eighteen Finnish university students of English, divided into control and experimental groups, provided speech samples at the beginning and end of an L2 English pronunciation course. The samples were analyzed with a mixed-methods approach. Quantitative analyses demonstrated mostly minor differences in fluency development across the groups. Qualitative analyses demonstrated individual learning trajectories regarding stalling mechanisms. The findings suggest that while corpus-based awareness-raising activities can be beneficial for supporting L2 fluency development, a more extensive teaching approach is needed for full effects. The study has implications for L2 fluency research and pedagogy.

**Keywords:** fluency, intervention, pedagogy, second language speech, mixed-methods

## 1. Introduction

Fluency is a key aspect of second language (L2) oral proficiency, which has been extensively studied from both learner corpus and SLA perspectives (see, e.g., Degand et al., 2019; Tavakoli & Wright, 2020). While several studies characterize L2 speech fluency based on analyses of fluency-related features using learner corpus data or other L2 speech data, pedagogically oriented studies on the effects of fluency instruction on L2 fluency development in formal instruction are rare (but see Tavakoli et al., 2016). The effects of teaching on fluency development appear somewhat mixed based on previous results, potentially due to the various teaching approaches that have often focused on the development of general speaking skills rather than speech fluency specifically and point to the need for further research in the field. A potential novel approach for the effective teaching of L2 speech fluency involves the use of authentic examples of learner speech for awareness-raising purposes. While several L2 fluency researchers have made recommendations for effective fluency teaching practices, including the use of learner corpora (e.g., Derwing, 2017; Lintunen et al., 2020; Peltonen, 2020), there is a lack of empirical studies putting these pedagogical recommendations into practice and examining their efficiency in facilitating L2 speech fluency development.

To address the two gaps in L2 speech fluency research – research on the effects that fluency instruction has on fluency development in formal instruction and the efficiency of corpus-based teaching materials on facilitating fluency development – the present study explores the influence of a new teaching approach utilizing learner-corpus-based awareness-raising activities on the development of L2 fluency. In the study, 18 Finnish university students of English participated in a mandatory course that aimed to improve

their L2 English pronunciation, including suprasegmental aspects related to L2 fluency. The course consisted of twelve 90-minute language laboratory sessions once a week. The students were enrolled in one of the two groups taught in parallel by the same teacher, and these teaching groups were assigned the status of an experimental group ($n$ = 10) and a control group ($n$ = 8). Both groups received English pronunciation and fluency teaching based on listen-and-repeat and reading aloud exercises, while the experimental group also received additional instruction specifically aimed at raising awareness of fluency-related features. The approach targeted three aspects that have been identified in previous research as key elements in the teaching of L2 fluency: the role of L1 individual speaking style in L2 speech fluency, fluency-enhancing strategies, and formulaic sequences (e.g., Lintunen et al., 2020; Peltonen, 2020; Tavakoli et al., 2016).

Based on a mixed-methods approach (see, e.g., Degand et al., 2019; Peltonen, 2020; Peltonen & Lintunen 2016), the present study addressed two research questions (RQs):

RQ 1) What kinds of differences can be observed in L2 fluency performance between pre- and post-tests for the control and experimental group?

RQ 2) What does variation in the use of stalling mechanisms (filled pauses [FPs] and repetitions) across individual participants reveal about the differences between the control and experimental groups in their development?

The quantitative analyses focused on the potential differences in fluency development across the control and experimental groups (RQ 1). To illustrate individual learning trajectories and individual differences among learners across the groups, a complementary qualitative analysis of two fluency-enhancing strategies or *stalling mechanisms* (FPs and repetitions; e.g., Dörnyei & Kormos, 1998) was conducted (RQ 2).

The article is structured as follows: in Section 2, the theoretical background, with focus on previous research on L2 speech fluency development, is presented. In Section 3, the methodology and materials of the present study are discussed, starting with a description of the participants (Section 3.1), followed by the data collection (Section 3.2) and the procedure and data analysis (Section 3.3). The results are presented in Section 4, followed by a discussion of the findings in Section 5. Concluding remarks are made in Section 6.

## 2. Theoretical background

Fluency is an essential aspect of L2 oral proficiency that can be examined separately from other dimensions of proficiency, such as accuracy or complexity (on the complexity-accuracy-fluency [CAF] framework; see, e.g., Housen et al., 2012). In contrast to the so-called *broad sense of fluency* as overall (oral) proficiency (Lennon, 1990), in SLA and learner corpus research, *a narrow sense of fluency* is often applied (ibid.). In this sense, fluency refers to one element of a broader oral competence and is captured with the notions of smoothness and effortlessness. To analyze these elements of L2 speech, various temporal measurements related to the speed of talk, the extent of pausing, and the rate of corrections are commonly used in fluency studies. Skehan's (2009) framework of speed fluency (e.g., articulation rate), breakdown fluency (e.g., different aspects of pausing), and repair fluency (e.g., self-corrections, repetitions) has become well established in the field to examine utterance fluency (Segalowitz, 2010)

comprehensively; that is, for measuring fluency-related aspects from L2 speech samples that create an impression of fluent speech (perceived fluency; Segalowitz, 2010).

The present study focuses on fluency in the narrow sense (utterance fluency) based on Skehan's (2009) three-fold framework and complements the analysis with measures of so-called fluency resources or stalling mechanisms (such as FPs, repetitions, or filler words; e.g., Dörnyei & Kormos, 1998) to examine whether the participants utilize individual mechanisms to avoid long silences during planning for maintaining fluent speech. These features have been identified as being particularly prone to individual variation in previous L2 speech fluency research (e.g., Cucchiarini et al., 2002; Dumont, 2018; Götz, 2013; Peltonen, 2021; Tavakoli et al., 2020) and have been shown to co-occur (e.g., Peltonen & Lintunen, 2016). In addition, some composite measures combining Skehan's (2009) speed and breakdown dimensions were used to obtain a more general impression of fluency development and to complement the measures related to individual dimensions of fluency.

The different fluency measures have been widely used to study learners' oral proficiency development. While several cross-sectional studies have documented increases regarding at least some aspects of fluency when the proficiency level improves (e.g., Kormos & Dénes, 2004; Tavakoli et al., 2020), especially long-term longitudinal investigations tracking fluency development have been rare (but see Derwing et al., 2008; Hanzawa, 2021). However, studies examining the effects of study abroad (SA) on fluency development form a notable exception to the lack of longitudinal L2 fluency studies (e.g., Huensch & Tracy-Ventura, 2017; Lennon, 1990; Mora & Valls-Ferrer, 2012; Segalowitz & Freed, 2004; Tavakoli, 2018; Towell, 2002; Towell et al., 1996). Generally speaking, while studies on SA have documented gains in fluency development especially for

composite measures, such as speech rate (e.g., Tavakoli, 2018), fluency development may also vary across different dimensions of fluency (e.g., Huensch & Tracy-Ventura, 2017) and several background factors, notably the use of the target language during SA (e.g., Mora & Valls-Ferrer, 2012), may affect the development. It is also noteworthy that studies in the SA context are often concerned with relatively long-term developmental patterns (ranging from several months to even years; Borràs & Llanes, 2021), while fluency studies examining development in other contexts have typically involved comparably shorter investigations.

While relatively rare, some studies have examined L2 fluency development in formal teaching contexts (Peltonen, 2021; Peltonen & Lintunen, 2019), sometimes along with accuracy and complexity (Doe, 2021; Tavakoli et al., 2016; Tonkyn, 2012). Of these studies, Tavakoli et al.'s (2016) study exploring the impact of a fluency-focused pedagogical approach on fluency development comes closest to the present study in design, while others have examined the impact of regular fluency activities (Doe, 2021) or the effects of general speaking skills instruction on fluency development (Peltonen, 2021; Peltonen & Lintunen, 2019; Tonkyn, 2012). Similarly to the present study, the studies have, overall, been concerned with exploring short-term effects, ranging from a four-week intervention (Tavakoli et al., 2016) to a 14-week semester (Doe, 2021). The results of the studies provide somewhat mixed results, potentially due to the differences in target groups, duration of the studies, and the teaching approaches.

Of the studies examining general speaking skills instruction, Peltonen and Lintunen (2019) and Peltonen (2021) examined Finnish upper secondary school learners' (17–18-year-olds) fluency development during a 7-week course on spoken English. Peltonen and Lintunen's (2019) study ($N$ = 20) based on monologue speech fluency tasks (reporting an

L1 Finnish news item in the L2) demonstrated statistically significant fluency gains for two composite (speech rate and phonation-time ratio) and one breakdown fluency measure (mean length of silent pauses, SPs), but not for the three other measures included in the study (articulation rate, mean length of run, or the number of SPs). The results also showed individual differences in the participants' developmental patterns and differences across the two instructional groups. Peltonen's (2021) study with a similar population ($N$ = 10) included peer interactional data (a discussion task), enabling the study of both individual and interactional fluency development. Based on 13 measures of individual fluency and four measures of interactional fluency included in the study, only the average turn length showed a tendency towards improved fluency based on group averages, but the difference did not reach statistical significance. Similarly to Peltonen and Lintunen (2019), substantial within-group variation in the development was also observed on several fluency measures in the study.

In contrast to the upper secondary school participants examined in Peltonen and Lintunen (2019) and Peltonen (2021), Tonkyn (2012) analyzed the development of spoken L2 English among adult university-level students in the UK. The study examined CAF development during a 10-week English for academic purposes (EAP) course ($N$ = 24; various L1s) based on interview data. Based on the nine fluency measures included in the study, mainly capturing the speed and breakdown fluency dimensions, the overall development in fluency was found to be relatively limited, as statistically significant improvement was found only for the measures of mean length of fluent runs and mean turn length. Discussing the findings, Tonkyn (2012: 239) suggests that there might be a potential trade-off between fluency and complexity, in line with Skehan's (2009) trade-off hypothesis.

While the three studies examining short-term fluency development based on general speaking skills instruction demonstrated relatively modest gains in L2 fluency, somewhat clearer evidence for the benefits of short-term instruction on L2 speech fluency development has been provided in a study by Tavakoli et al. (2016). In their teaching experiment, EAP students in the UK (with various L1s) participated in general speaking and listening activities (control group, *n* = 18) or fluency-focused instruction incorporating additional awareness-raising activities and fluency strategy training (experimental group, *n* = 19) over a four-week period. Based on an analysis of CAF indices, statistically significant improvement in fluency for the experimental group was found for four out of nine measures (speech rate, articulation rate, mean length of run, and phonation-time ratio), whereas the control group did not demonstrate fluency gains.

Another recent study incorporating targeted fluency practice (3/2/1 speaking activities; a variation of the 4/3/2[1] activity, see De Jong & Perfetti, 2011) was conducted by Doe (2021) among 32 Japanese university students enrolled in L2 English discussion classes (four classes of varying levels; no control group). Tracking the development of CAF over 14 weeks, the findings showed small but statistically significant gains for two (mean length of pauses and phonation-time ratio) of the four measures of fluency included in the study. Different proficiency level classes also showed somewhat differing patterns, clearest gains being obtained in the highest proficiency group (high intermediate) for both measures. Furthermore, within-group variation was large for all four classes.

Based on these five studies exploring fluency development in formal education, some short-term improvement in fluency seems to be possible even without targeted

[1] In this activity, the participants first perform the task by speaking (e.g., telling a story) for four minutes, followed by conveying the same content in three minutes to a different listener, and finally exchanging partners once more to explain the contents in two minutes.

fluency instruction, especially regarding length-based (turn length) measures or composite fluency measures combining elements of speed and pausing. However, the results regarding gains for specific dimensions of fluency (speed, breakdown, and repair) are less clear. So far, the most compelling evidence for short-term fluency development comes from Tavakoli et al. (2016), who demonstrated development in a short time (four weeks) for composite and speed fluency measures for the experimental group, pointing to the benefits of targeted fluency instruction. Doe's (2021) results also suggest that focused fluency practice can be beneficial, despite his study showing more limited gains compared to Tavakoli et al. (2016). Furthermore, as Doe's (2021) study did not include a control group, it is less clear whether the observed improvement can be attributed to the regular fluency practice or some other factor. To fill the research gap regarding the short-term effects of fluency instruction on L2 fluency development in formal teaching using corpus-based approaches, the present study explores the effects of fluency teaching in the context of Finnish university students of English participating in an L2 English pronunciation course, with the experimental group receiving additional fluency instruction (within the same number of teaching hours) using learner-corpus-based awareness-raising activities.

Since previous studies exploring the impact of corpus-based fluency teaching activities are rare, a set of learner-corpus-based awareness-raising activities was developed for the teaching experiment of the present study. Our experiment targeted three key areas identified as potentially beneficial for fluency development (e.g., Lintunen et al., 2020; Peltonen, 2020; Tavakoli et al., 2016): the connections between L1 individual speaking style and L2 speech fluency, fluency-enhancing strategies (e.g., Götz, 2013), and formulaic sequences. The development of the activities was based on previous

research on fluency activities (e.g., Rossiter et al., 2010), and the activities were centered around the idea of raising the participants' awareness of fluency-related features (for a similar approach, see, e.g., Tavakoli et al., 2016). Especially regarding the use of stalling mechanisms, awareness-raising activities have been suggested as being particularly helpful in highlighting their potential role as fluency-enhancing resources (see Dörnyei, 1995; Guillot, 1999; Peltonen, 2020), but we extended this approach to L1 speaking style and formulaic sequences. L2 fluency studies have drawn attention to the role of L1 speaking style in influencing L2 speech fluency (for an overview, see Peltonen, 2020), but to our knowledge, this has not yet been included in fluency teaching approaches. Yet, awareness of one's own speaking style and its effects on L2 speech fluency may potentially reduce anxiety related to L2 speaking and facilitate individual goal setting for L2 learners (Lintunen et al., 2020: 198). Furthermore, formulaic sequences are likely to be useful targets of instruction, as they can facilitate fluent speech production from a processing perspective (e.g., Wood, 2006). Overall, since language teachers tend to define fluency relatively broadly and mostly promote fluency in classrooms based on general speaking activities rather than fluency-focused activities (Tavakoli & Hunter, 2018), by developing research-based materials specifically designed to enhance learners' awareness of speech fluency, the present study contributes not only to the development of pedagogically oriented L2 fluency research, but also to L2 fluency teaching practices.

## 3. Methodology and materials

### 3.1 Participants

The participants in the present study were 18 Finnish university students of English (16 female, 2 male). The mean age of the participants was 27 years (median 25 years), and they had studied English for an average of 10.2 years at school. All participants were students of English at university; most participants ($n$ = 17) studied English as their minor subject and had studied English at university for an average of 0.75 years. All participants spoke Finnish as their first language, and one participant also reported speaking English at home (Participant 1 in the control group). No official proficiency level test of English was administered to the participants, but based on the target level of B2 in the Common European Framework of Reference for Languages (CEFR; Council of Europe, 2001) for an average student at the end of upper secondary school (Finnish National Agency of Education, 2019: 177) and the participants having passed a competitive entrance exam to study English at university, they were estimated to represent roughly the level C1/C2 in the CEFR. All students participated in the study voluntarily, and signed informed consent was obtained from them. The participants were also informed that the data collection was part of a larger project on L2 speech fluency (*Fluency and Disfluency Features in L2 Speech*, funded by the Research Council of Finland 2020–2024; decision number 331903).

### 3.2 Data collection

The data for the present study were collected during a BA level, compulsory course on L2 English pronunciation included in the basic studies for major and minor students of English. Based on the description of the course offered in the study guide, the course

aimed to improve the students' pronunciation, rhythm, and fluency in English. In practice, generally the emphasis on the course is on segmental aspects of English that are difficult, in particular, for Finnish learners, while suprasegmental aspects are covered to a lesser extent (for a previous pronunciation-related study conducted in a similar context, see Lintunen & Mäkilähde, 2018). Along with practical speaking tasks, such as listen-and-repeat and read-aloud exercises aimed at improving the students' L2 English pronunciation, the course includes phonetic transcription exercises. The participants attended twelve 90-minute language laboratory sessions once a week during spring 2022 in one of the two groups taught in parallel, and these teaching groups were assigned the status of a control group ($n$ = 8) and an experimental ($n$ = 10) group.[2] Both groups were taught by the first author, and the total amount of teaching received was the same for both groups (18 hours). Pre- and post-tests were administered at the beginning of the course (January 2022; T1) and at the end of the course (April 2022; T2). The tests involved picture description tasks and read-aloud tasks, but only the data from the picture description tasks were included in the present study, as the focus was on examining the development of L2 speech fluency in elicited but free production.

Two cartoon strip prompts were used for the picture description task. To control for potential order effects, the tasks were counterbalanced across the groups. To prepare for the task, the participants were given two minutes of planning time, and they were instructed to describe the pictures and to tell the story in their own words at their own pace in English. The participants were allowed to look at the cartoon during the task. Picture descriptions have been commonly used to elicit data in fluency studies (e.g.,

[2] Of the original 20 participants, one participant did not take the post-test and another participant was excluded from the sample due to producing clearly shorter samples compared to the other participants (below 100 syllables).

Lennon, 1990), and the cartoon strips used in the present study have been used in previous studies (e.g., Peltonen, 2020). Both strips included six frames, a clear storyline, and only minimal text (one thought bubble).

The teaching experiment applied in the present study included targeted fluency activities for the experimental group. The control group did not receive this targeted fluency practice but received the same amount of teaching. Thus, while both groups received English pronunciation and fluency teaching based on listen-and-repeat and reading aloud exercises during the course, the experimental group received instruction focusing on three fluency features during two of the language laboratory sessions: the role of L1 individual speaking style in L2 speech fluency, fluency-enhancing strategies, and formulaic sequences. As mentioned in Section 2, these features were chosen due to them being highlighted in previous research as important aspects of L2 speech fluency and potentially useful as the basis for L2 fluency teaching. The instructional approach, focused on awareness-raising of these features, was based on three steps, involving 1) teacher-led introduction to the target feature, 2) exposure to the feature via authentic speech samples, and 3) practice in identifying the feature from one's own speech (see also Tavakoli et al., 2016). The novelty of the present study was that examples from learner corpora were used in step 2 to facilitate the participants' awareness-raising regarding individual speaking styles, fluency-enhancing strategies, and formulaic sequences with the help of authentic L1 and L2 speech samples rather than by using invented examples or only the participants' own speech. The learner corpus used in the teaching experiment was compiled at the Department of English, University of Turku, and included both monologue L2 English speech samples (picture description) as well as corresponding L1 Finnish samples of the monologues from the same learners (see Peltonen, 2020). A

practice task (including steps 2 and 3) incorporating individual speaking style and fluency-enhancing strategies is illustrated in Appendix 1.

### 3.3 Procedure and data analysis

All speech samples were transcribed by the first author. Seven fluency measures (see Table 1) were chosen based on previous research (see De Jong, 2018; Kormos, 2006; Peltonen, 2020) to reflect the three main dimensions of fluency (speed, breakdown, and repair; e.g., Skehan, 2009). Furthermore, two composite measures combining the speed and breakdown dimensions were included. In selecting the fluency measures, comparability with previous fluency development studies in the formal teaching context and the characteristics of the teaching approach of the present study were considered. The first two target areas of the teaching experiment, individual speaking style and fluency-enhancing strategies, were captured with two measures: the number of FPs per minute and the number of repetitions per minute. While these aspects have traditionally been considered indicators of disfluency, they were incorporated in the present study as potential stalling mechanisms that can facilitate coping with processing time pressure, reduce time spent in silence and thus contribute to maintaining fluent speech production during difficulties (e.g., Dörnyei & Kormos, 1998; Götz, 2013; Peltonen, 2020; Wolk et al., 2020). The third target area of the experiment, formulaic sequences, is linked to the mean length of utterance measure (comparable to mean length of run), since the use of formulaic sequences can contribute to an increased utterance length (e.g., Wood, 2006). All fluency measures with their operationalizations are included in Table 1. The frequency measures were standardized per minute of speaking time (De Jong, 2016).

**Table 1**. Fluency measures and their operationalizations

| **Fluency dimension** | **Fluency measure** | **Operationalization** |
|---|---|---|
| Speed | 1. Articulation rate | Syllables per minute of speaking time |
| Breakdown | 2. Number of SPs per minute | Silences above 0.25 seconds in duration |
| | 3. Mean duration of SPs | Total pause time / number of SPs |
| | 4. Number of FPs per minute | Non-lexicalized FPs, e.g., *uh, um* |
| Repair | 5. Number of repetitions per minute | The number of words or longer stretches of the participant's own speech that were repeated without modification |
| Composite | 6. Speech rate | Syllables per minute of total time |
| | 7. Mean length of utterance | Total speaking time / number of utterances (utterance = stretch of speech between SPs) |

To prepare the data for fluency analysis, the syllables were counted manually by the first author (measures 1, 6). For measuring breakdown fluency (measures 2–4), SPs were identified from the data. The minimum duration for a SP was set at 0.25 seconds; shorter pauses were treated as micropauses and were not included in SP analyses. SP times exceeding 3 seconds were excluded from SP duration and sample total duration

calculations to prevent them from skewing the distributions of the speech rate and mean SP duration variables, as such lengthy silences are not likely to reflect individual disfluencies. The SPs were identified with a two-step procedure: first, a Praat script (De Jong & Wempe, 2009) was used. Second, the SP boundaries were checked manually and adjusted. For measures 4 and 5, FPs and repetitions (together referred to as stalling mechanisms) were manually identified from the data. After the data preparation, to calculate the fluency measures, the durations and frequencies of the Praat annotations were extracted with a script (Lennes, 2002). The average duration of the speech samples in the control group at T1 was 107.15 seconds and 81.10 seconds for the experimental group. At T2, the average duration of the samples was 70.13 seconds for the control group and 97.84 seconds for the experimental group.[3] The average number of syllables changed from an average of 246 to 178 for the control group from T1 to T2, while the experimental group produced, on average, 193 syllables at T1 and 223 syllables at T2.

The statistical analyses used for answering RQ 1 consisted of two-way mixed between/within ANOVAs (see, e.g., Larson-Hall, 2010). The dependent variables were the seven L2 fluency measures, and the independent variables were group (two levels: experimental group and control group) and time (two levels: pre-test and post-test), the former being a between-subjects variable and the latter a within-subjects variable. Based on Shapiro-Wilk's tests of normality, visual examinations of boxplots, and Levene's tests, the assumptions for repeated-measures ANOVA were sufficiently met.[4] Due to multiple comparisons, false discovery rate (FDR) correction was applied to the p-values of the

[3] Due to a technical malfunction, three of the experimental group's T2 recordings were cut short. However, the recordings were of sufficient length for reliable analyses, and the effect of the malfunction on the results is reduced due to the standardization of all fluency measures.

[4] According to Shapiro-Wilk's tests, only six of the total 28 examined variables (fluency measures by time and group) were non-normally distributed. Levene's tests indicated homogeneity of variances for all variables ($p > .05$).

interaction effects to control for type I error, i.e., potential false positives (see Larson-Hall, 2010: 251–252). It should be noted that this procedure, coupled with the relatively small, intact teaching groups used in the present study, poses limitations for statistical power, and can be reflected in a heightened probability for a type II error, i.e., potential false negatives. To counteract the potential limitations of the small sample, the present study utilized a mixed-methods approach and incorporated a qualitative analysis focused on individual differences and learning trajectories in the use of the two stalling mechanisms (RQ 2). By highlighting individual participants' developmental patterns and within-group variation, the analysis complements the information provided by the quantitative analyses focusing on group-level tendencies (on mixed-methods analyses in L2 speech fluency research, see, e.g., Peltonen, 2020).

## 4. Results

Before answering RQ 1 based on the mixed ANOVAs (Table 3), we discuss the descriptive statistics (M, SD) for the control and experimental groups based on Table 2 (for individual values for each participant, see Appendix 2; for raw frequencies, see Appendix 3). Variation in fluency development across individual speakers regarding the use of stalling mechanisms (RQ 2) is discussed based on Figures 1 and 2 and examples from the data.

**Table 2.** Fluency measures (M, SD) for the control group (CG) and experimental group (EG) at T1 and T2

| Measure | CG T1<br>M (*SD*)<br>(*n* = 8) | EG T1<br>M (*SD*)<br>(*n* = 10) | CG T2<br>M (*SD*)<br>(*n* = 8) | EG T2<br>M (*SD*)<br>(*n* = 10) |
|---|---|---|---|---|
| 1. Articulation rate (syll. / min.) | 222.39 (*21.71*) | 206.78 (*24.98*) | 220.26 (*21.64*) | 204.86 (*24.19*) |
| 2. Number of SPs / min. | 40.97 (*7.58*) | 34.94 (*9.69*) | 36.42 (*11.24*) | 35.86 (*6.02*) |
| 3. Mean length of SPs (sec.) | 0.89 (*0.16*) | 0.78 (*0.09*) | 0.77 (*0.11*) | 0.85 (*0.12*) |
| 4. Number of FPs / min. | 6.23 (*5.79*) | 9.43 (*6.66*) | 4.39 (*3.86*) | 7.84 (*5.33*) |
| 5. Number of repetitions / min. | 2.34 (*3.83*) | 2.44 (*2.37*) | 1.40 (*1.67*) | 3.83 (*3.17*) |
| 6. Speech rate (syll. / min.) | 139.92 (*21.05*) | 142.26 (*17.00*) | 152.20 (*21.77*) | 136.35 (*17.50*) |
| 7. Mean length of utterance | 5.48 (*1.22*) | 6.02 (*1.31*) | 6.49 (*2.57*) | 5.72 (*1.24*) |

Note: T1 = Pre-test (beginning of course), T2 = Post-test (end of course), M = mean, SD = standard deviation.

Based on the descriptive statistics in Table 2, the changes in fluency from T1 to T2 were relatively minor for both the control group (CG) and the experimental group (EG). Based on the group averages, there are some trends towards increased fluency between T1 to T2

for the CG: the number of SPs per minute was reduced and the average length of SPs decreased, coupled with a somewhat higher speech rate and longer utterances. However, articulation rate was slightly slower at T2 on average for the CG. Regarding the averages for the two stalling mechanisms, both FPs and repetitions were used, overall, somewhat less frequently at T2 compared to T1 in the CG. However, as indicated by the standard deviations, there was substantial within-group variation in the use of these mechanisms.

The EG showed a slightly different trend based on the descriptive statistics, as the group averages for articulation rate, speech rate, mean length of utterance, and both SP measures indicated a slightly less fluent performance at T2 compared to T1. Yet, it should be noted that the starting point at T1, on average, was somewhat lower for the CG compared to the EG based on the fluency measures. Regarding the two stalling mechanisms, the patterns for FPs were relatively similar to the CG: the EG participants also used somewhat fewer FPs, on average, at T2 compared to T1. However, compared to the CG that showed a tendency towards fewer repetitions at T2 compared to T1, the pattern for the EG was the opposite: the EG participants used repetitions, on average, somewhat more frequently at T2 compared to T1. It is also noteworthy that as with the CG, the within-group variation was substantial regarding both stalling mechanisms at T1 and T2, as indicated by the high standard deviations.

**Table 3.** Results of the mixed between/within ANOVAs with group (experimental vs. control) and time (T1 vs. T2) as factors

| ***df* (1, 16)** | **Group** | | | **Time** | | | **Interaction** | | |
|---|---|---|---|---|---|---|---|---|---|
| **Measure** | ***F*** | ***p*** | $\eta^2_p$ | ***F*** | ***p*** | $\eta^2_p$ | ***F*** | ***p*** | $\eta^2_p$ |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1. Articulation rate | 2.154 | .162 | .119 | .367 | .553 | .022 | .001 | .975 | .000 |
| 2. Number of SPs / min. | .728 | .406 | .044 | .953 | .344 | .056 | 3.109 | .097 | .163 |
| 3. Mean length of SPs | .082 | .778 | .005 | .247 | .626 | .015 | 7.476 | .015 | .318 |
| 4. Number of FPs / min. | 2.006 | .176 | .111 | 1.981 | .178 | .110 | .010 | .920 | .001 |
| 5. Number of repetitions / min. | 1.075 | .315 | .063 | .349 | .563 | .021 | 3.722 | .072 | .189 |
| 6. Speech rate | .609 | .447 | .037 | .605 | .448 | .036 | **10.461** | **.005** | **.395** |
| 7. Mean length of utterance | .026 | .873 | .002 | .857 | .368 | .051 | 4.594 | .048 | .223 |

Note: Statistically significant results (after the FDR correction) are highlighted in bold.

$\eta^2_p$ = partial eta squared.

Despite some trends indicating slight changes in fluency for both the CG and EG based on the descriptive statistics, the mixed between/within ANOVAs, overall, did not indicate statistically significant differences between the groups or over time (see the results for the main effects of Group and Time in Table 3). However, as our main interest was on the potential differences in development between the CG and EG, the results regarding the interaction effects answer our RQ 1 more directly. As indicated in Table 3, some interactions reached statistical significance ($p < .05$), indicating that the developmental patterns from T1 to T2 differed between the two groups for the variables mean length of SPs, speech rate, and mean length of utterance; that is, for one breakdown fluency

measure and the two composite measures. However, after the FDR correction was applied to the *p*-values regarding the interactions, only the interaction for speech rate remained statistically significant. In other words, the difference observed based on descriptive statistics, namely the CG increasing their speech rate from T1 to T2 while the pattern was the opposite for the EG, was confirmed to be significant with the ANOVA (with a large effect size).

To complement the analyses of group-level tendencies, the individual variation in fluency was explored from the perspective of the two stalling mechanisms, FPs and repetitions. The distributions of FPs and repetitions are illustrated in Figures 1 and 2, respectively.

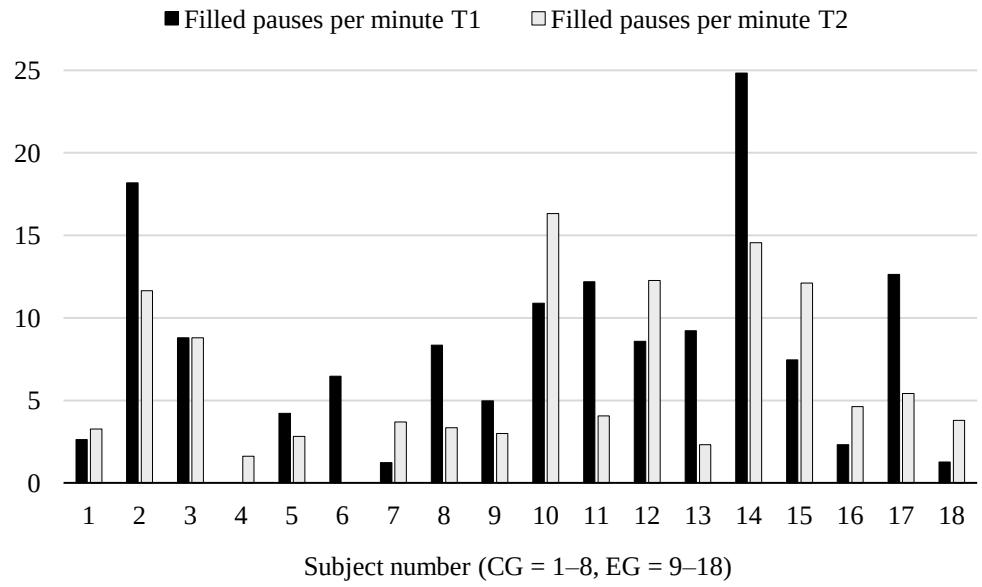


**Figure 1.** Distribution of FPs for the whole sample (*N* = 18)

As can be seen in Figure 1, the individual differences in the use of FPs were large for both groups (CG T1 range 0–18.9, EG T1 range 1.26–24.83; CG T2 range 0–11.64, EG T2

range 2.31–16.33). Regarding the individual developmental patterns, in the CG, an increase in the use of FPs was observed for three participants (1, 4, 7), while four participants (2, 5, 6, 8) used fewer FPs at T2. The number of FPs stayed the same for one participant (3). Despite the group averages pointing to a tendency of fewer FPs in both groups, in fact, the distribution between decrease and increase in the use of FPs was even more equally divided in the EG compared to the CG: half of the group, i.e., five participants (10, 12, 14, 15, 16, 18) used more FPs at T2 compared to T1, while half used fewer FPs at T2 (participants 9, 11, 13, 14, 17).

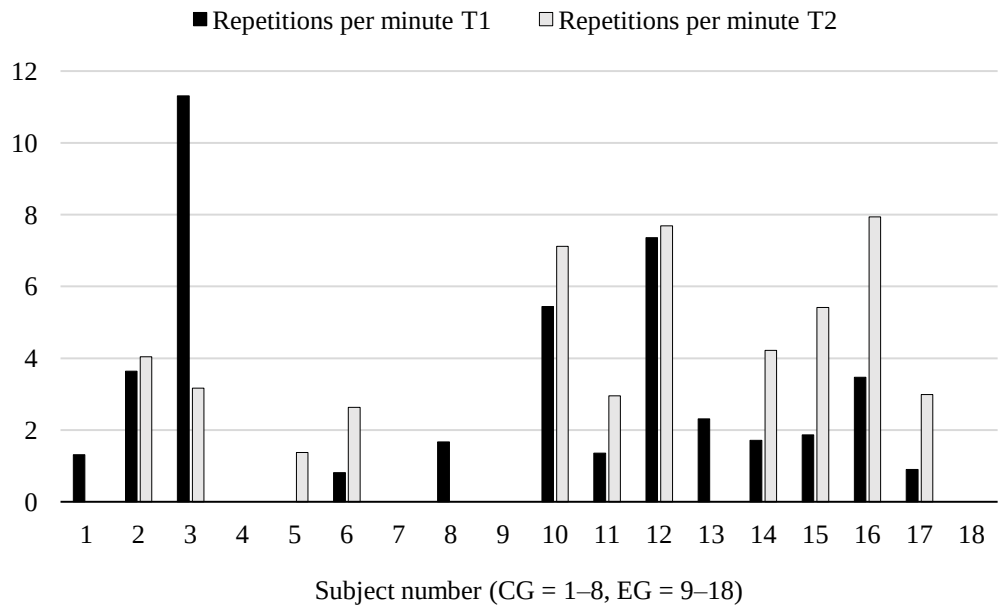


**Figure 2.** Distribution of repetitions for the whole sample (*N* = 18)

As illustrated in Figure 2, even more variation was observed in the use of repetitions compared to FPs (CG T1 range 0–11.31, EG T1 range 0–7.36; CG T2 range 0–4.04, EG T2 range 0–7.94). In the CG, where the group mean suggested a slight reduction in the use of repetitions, three participants used fewer repetitions at T2 compared to T1 (1, 3,

8), while three participants (2, 5, 6) increased their use of repetitions. Additionally, two participants (4, 7) did not produce any repetitions at T1 or T2. Compared to the CG, the EG participants displayed a somewhat more homogeneous pattern regarding the use of repetitions. The group mean indicated a tendency towards an increase in the use of repetitions, and for seven participants, some increase was observed (10, 11, 12, 14, 15, 16, 17). A decrease was observed in the use of repetitions for only one participant (13), while two participants (9, 18) did not use repetitions at all either at T1 or T2. Thus, for the EG, compared to the repetition patterns observed for the CG or the relatively mixed profile of FPs detected in both groups, the tendency was towards an increase in the use of repetitions.

While the increase in repetitions in the EG was relatively subtle for some (e.g., participant 12), there were also participants who more than doubled their use of repetitions (e.g., 14, 15, 16, 17). For instance, participant 15 used 1.86 repetitions per minute at T1, but 5.41 repetitions per minute at T2. At T2, the participant seemed to use the repetitions as a stalling device to obtain more planning time, as the repetitions frequently co-occurred with SPs and/or FPs (Example 1b), while similar use was observed only once at T1 (Example 1a).

and starts to walk towards the (0.58) um (0.35) <u>the</u> snowman?

Example 1a (Participant 15, T1)

so they are planning to cut it down an (0.29) building a home? (0.79) <u>home</u> for themselves

Example 1b (Participant 15, T2)[5]

This type of use of repetitions as stalling devices was detected among other participants in the EG as well, especially at T2, while it was rarer in the CG. Interestingly, among the seven participants in the EG increasing their use of repetitions, three reduced their use of FPs from T1 to T2 (11, 14, 17). For these participants, this pattern could indicate a shift from one stalling strategy to another. For instance, participant 14 used FPs at T1 most frequently in the whole sample but decreased their use by T2. At the same time, they increased their use of repetitions from T1 to T2. As illustrated in Example 2b, the use of repetitions as stalling mechanisms in conjunction with SPs and/or FPs, suggesting that they helped in reducing time spent in silence during planning, became more prominent at T2, while this was rare at T1 (Example 2a).

the kid has (.) uh to put (.) <u>put</u> it on a snowman

Example 2a (Participant 14, T1)

the tree is now (0.35) big and uh (0.60) <u>and</u> fully grown?

Example 2b (Participant 14, T2)

Among the other four participants in the EG who increased their use of repetition, an increase in FPs was also observed. For these participants, overall increase in the use of stalling mechanisms was thus detected from T1 to T2. Of these participants, especially

[5] In the examples, the question mark indicates rising intonation, and SP durations are given in parentheses. Micropauses of less than 0.25 seconds are marked with "(.)".

the samples produced by participants 10 and 16 provide some evidence towards the use of these stalling devices not only becoming more frequent, but also more efficient over time. Both participants had some sequences in their T1 samples where they seemed to "get stuck" and used stalling devices, but still produced lengthy SPs of around 2 seconds (Example 3a). In contrast, at T2, in similar situations, stalling mechanisms were used by the participants more efficiently, resulting in shorter SPs (Example 3b).

> and the (0.71) hat ends up (0.27) on (1.51) on a(hh) uh (2.59) uh (.) not (0.29) on the snowman but
>
> Example 3a (Participant 10, T1)

> the (.) tree has been cut down there's no more (0.90) no more tree there and uh (0.64) uh I think (0.86) a lot of time has passed by
>
> Example 3b (Participant 10, T2)

In other words, the participants seemed to become more skilled in using stalling mechanisms to get over speech production problems: they managed to reduce SP time and keep the flow of talk going during planning.

## 5. Discussion

The first research question addressed the differences in L2 fluency performance across T1 and T2 for the CG and EG. Based on a series of two-way mixed between/within

ANOVAs, few statistically significant differences in the developmental patterns between T1 and T2 for the two groups were observed. In other words, overall, the fluency-focused teaching given for the EG did not result in greater group-level fluency development compared to the CG based on the quantitative analyses. However, the statistically significant interaction between group and time for speech rate suggests that the developmental patterns differed on this particular measure, albeit in a somewhat unexpected direction: the CG produced faster speech at T2 compared to T1 based on the group average, potentially due to a somewhat lower speech rate at T1 compared to the EG, while the EG demonstrated the opposite pattern, producing slower speech at T2 compared to T1. These findings could be indicative of increased self-awareness among the EG due to the teaching approach: as these participants became more aware of details in their speech through the awareness-raising approach, they could have been monitoring the features of their speech fluency more carefully than the CG participants, resulting in somewhat less fluent speech, on average. It is possible that this effect is a temporary one, but to confirm this empirically, future research should incorporate a delayed post-test in the data collection design.

Overall, the quantitative results answering RQ 1 complement previous longitudinal L2 fluency studies tracking fluency development during SA, and, more specifically, in formal settings over a relatively short period of time (typically ranging from four to 14 weeks). Compared to the bulk of SA fluency studies, which generally point to gains in L2 fluency in environments where the target language is spoken and after relatively long-term exposure (ranging from several months to even years), the results of the present study point to less prominent changes over time. However, in the context of fluency studies focusing on formal education and short-term gains in fluency, the results are more

in line with previous findings that have shown relatively limited development (Peltonen, 2021; Peltonen & Lintunen, 2019; Tonkyn, 2012). Yet the finding of the CG increasing their speech rate (a composite measure) from T1 to T2 supports some previous findings of general speaking skills instruction facilitating fluency development especially when measured with composite or length-based measures of fluency (Peltonen, 2021; Peltonen & Lintunen, 2019; Tonkyn, 2012). Regarding the effects of the teaching approach, the overall relatively minor changes in L2 fluency observed in the present study are, nevertheless, in contrast with some previous findings involving explicit fluency instruction and/or regular fluency activities, which have demonstrated overall fluency improvement based on composite measures and for specific dimensions of fluency (breakdown fluency in Doe, 2021; speed fluency in Tavakoli et al., 2016).

Various factors can potentially explain the differing results compared to previous research. First, the participants were students of English at university, estimated to represent an advanced level of proficiency, while previous studies have examined university students of other subjects (Doe, 2021), students participating in an intensive EAP course (Tavakoli et al., 2016; Tonkyn, 2012), or younger learners at an intermediate level of proficiency in English (Peltonen, 2021; Peltonen & Lintunen, 2021). Due to the relatively high level of proficiency of the participants in the present study already at T1, there could have been less room for improvement in fluency compared to the populations examined in some previous fluency studies (on plateau effects in fluency development, see, e.g., Towell et al., 1996). Additionally, the relatively short duration of the course (12 weeks; 18 hours of teaching in total), coupled with its focus on segmental aspects of English pronunciation, could have influenced the findings; the time span of the experiment could have been too short and, especially for the CG, overall speaking skills

instruction not sufficient to facilitate substantial fluency development beyond the composite measure of speech rate (see also Tavakoli & Wright, 2020).

While Tavakoli et al. (2016) demonstrated fluency gains with their fluency teaching experiment in only four weeks, their teaching approach and the course schedule were more intensive compared to the present study. Tavakoli et al. (2016) had 8 sessions for fluency tasks and training for the experimental group (2 x 15–20-minute sessions per week) and the intensive EAP course (see also Tonkyn, 2012), overall, had 21 hours of teaching per week, while the groups in the present study met for 1.5 hours weekly and targeted fluency teaching and practice was only integrated into two sessions for the EG. Thus, it seems that a 'minimal' fluency teaching approach explored in the present study, including only 12 sessions altogether with the awareness-raising activities implemented in a focused manner on two sessions, was not extensive enough to fully facilitate group-level fluency development for the EG, compared to, for instance, the more intensive approach used by Tavakoli et al (2016). Furthermore, while we did not have access to the participants' L1 speech samples, it is also possible that especially those features that are potentially closely connected across L1 and L2 speech fluency (see, e.g., Peltonen, 2020) are even less prone to change in a short period of time than other fluency-related features. In future studies, this aspect could be examined empirically by collecting L1 speech samples from the participants along with L2 samples.

The findings regarding RQ 2 and individual variation in fluency development indicated that, at least for certain students, the teaching approach could have influenced some elements of fluency, especially stalling mechanisms, despite the lack of statistically significant group-level differences over time across the EG and CG regarding these features. Stalling mechanisms were found to be subject to within-group variation in both

groups, in line with previous findings (e.g., Cucchiarini et al., 2002; Dumont, 2018; Götz, 2013; Peltonen, 2021; Tavakoli et al., 2020). In addition, the qualitative analysis of this variation demonstrated that, especially in the EG, there was a tendency for several participants to increase their use of repetitions from T1 to T2. The findings suggest that some participants seemed to have internalized the possibility to use repetitions as stalling devices to avoid long silences during planning, as instructed during the teaching experiment. Furthermore, shifts in the types of stalling mechanisms and increases in the efficiency of stalling mechanism use were observed from T1 to T2 for some EG participants, further underscoring potential individual changes over time related to the teaching approach. Why a clearer increase was observed in repetitions rather than in the other type of stalling mechanisms, FPs, requires further study and could be related to preferences for particular stalling mechanism types (e.g., Dumont, 2018; Götz, 2013; Wolk et al., 2020). The location of FPs (clause external vs. clause internal pauses) could also be examined in the future to reveal potential further changes in their quality. The analysis of the two stalling mechanisms in the present study could also be complemented with analyses of other stalling mechanisms, such as drawls or filler words, to get a more comprehensive picture of the participants' individual stalling mechanism profiles. Finally, further analyses could target formulaic sequences specifically to reveal potential developmental patterns related to their use.

The present study offered insights into group-level and individual learning trajectories regarding L2 speech fluency based on a learner corpus approach to teaching fluency. As the participants of the present study were students of English at university and thus represented a relatively advanced proficiency level, future studies could implement similar approaches with learners at lower proficiency levels with potentially

more opportunity to increase their speech fluency. Comparisons with pseudo-longitudinal data from learners at different proficiency levels could also yield valuable results. Furthermore, the teaching materials could be adapted to various L2s and for different L1 groups. In fact, some of the activities have already been presented in a fluency-themed training for in-service language teachers that included participants teaching a range of different L2s (see Lintunen & Eirola, 2023; Peltonen & Lehtilä, 2023). Some of the teachers reported on their experiences in using these activities in their own classrooms (e.g., Mitrunen, 2023). While the teaching experiment in the present study covered three themes chosen from recent L2 fluency research, in the future, additional activities targeting other aspects of fluency could be designed. The activities included in the teaching experiment could also be developed further; for instance, they could include a final reflection component on the perceived usefulness of the tasks to further strengthen the awareness-raising approach used in the activities. With this type of reflection data, it could be possible to confirm potential other effects of the teaching approach beyond the changes in fluency features in the speech samples, such as changes in self-awareness, beliefs, and attitudes (see also Lintunen & Mäkilähde, 2018). Such an approach could also reveal whether there are individual preferences regarding this type of teaching, as it may be more suited to some learners than others.

## 6. Conclusion

The present study was among the first L2 speech fluency studies to explore the effects of teaching L2 speech fluency with learner-corpus-based awareness-raising activities. Based

on a quasi-experimental design, differences in the development of L2 speech fluency across the control and experimental groups over time were examined. While the quantitative analyses showed only minor differences between the two groups in L2 fluency development, qualitative analyses examining individual learners' use of stalling mechanisms suggested that the experiment potentially impacted at least some participants' use of repetitions. However, these results should be regarded as tentative due to the small sample size and preferably replicated in the future with larger sample sizes for more robust effects.

Nevertheless, the results suggest that the 'minimal' approach to teaching L2 fluency with a learner-corpus-based approach, at least among advanced learners of English in a formal teaching context, may not be sufficient to result in substantial fluency development in a short period of time and may, in fact, result in somewhat less fluent production, potentially due to increased self-awareness. Yet the observed shifts in stalling mechanism behavior for some participants over time suggest that the learner-corpus-based awareness-raising activities can have an impact on individual students' developmental patterns, highlighting the benefits of the mixed-methods designs in analyzing fluency development. Due to the nature of the course and its focus on segmental aspects of pronunciation mainly with listen-and-repeat and read-aloud exercises, in the future, a similar corpus-based L2 fluency teaching approach could be implemented on a general speaking skills course with more free production and discussion exercises to allow for the contents of the experiment to be integrated in the course contents more fully and perhaps resulting in more robust group-level changes. The results of the present study could also be compared to the students' development of other aspects of CAF, for example, pronunciation accuracy, in the future, to reveal potential trade-off effects (e.g.,

Skehan, 2009; see also Tonkyn, 2012). While left out of the scope of the present study, T1 and T2 recordings of the participants' read-aloud tasks could provide interesting points of comparison for the analyses of L2 fluency in freely produced speech in the future (see, e.g., Cucchiarini et al., 2002). The utterance fluency analyses of the present study could also be complemented with assessments of perceived fluency to examine whether any improvement can be detected based on holistic or analytic assessments of fluency.

Notwithstanding the limitations of the present study, the learner-corpus-based fluency teaching approach has the potential to provide the basis for future fluency teaching experiments in other contexts and with larger populations. The teaching materials developed for the present study can also help in-service teachers to incorporate fluency-focused practice in their L2 speaking skills instruction across various L1s and L2s. Due to the awareness-raising nature of the activities, however, teachers should bear in mind that some losses in fluency are possible: this point could be discussed in the classrooms together with the students so as not to discourage them if this happens as a result of increased awareness of fluency-related features and monitoring.

Overall, the study has contributed to the currently relatively small body of research examining short-term development of fluency in formal contexts; to complement these studies and long-term investigations in SA contexts, more research on long-term development of L2 fluency in formal education is needed. More specifically, the fluency teaching experiment approach used in the present study complements previous studies exploring L2 fluency pedagogy (Tavakoli et al., 2016) by using authentic speech samples to facilitate awareness-raising of fluency-related features. In particular, using samples from the same learners in their L1 and L2 provides a basis for exploring pedagogical approaches acknowledging the role of individual speaking style in L2 fluency, one of the

key themes in recent L2 fluency research. By utilizing previous corpus-based and SLA research on L2 speech fluency in designing the teaching experiment, including the example practice task presented in Appendix 1, the present study has helped to bridge the gap between these two fields of study and provides a first step in exploring how learner corpora and L2 fluency research together can inform L2 fluency pedagogy and the design of L2 fluency teaching materials in an ideal way.

## References


Borràs, J., & Llanes, À. (2021). Re-examining the impact of study abroad on L2 development: A critical overview. *The Language Learning Journal, 49*, 527–540.

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

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

De Jong, N. H. (2016). Fluency in second language assessment. In D. Tsagari & J. Banerjee (Eds.), *Handbook of second language assessment* (pp. 203–218). De Gruyter.

De Jong, N. H. (2018). Fluency in second language testing: Insights from different disciplines. *Language Assessment Quarterly, 15*, 237–254.

De Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods, 41*, 385–390.

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.

Degand, L., Gilquin, G., Meurant, & Simon, A. C. (2019). Introduction. In L. Degand, G. Gilquin, L. Meurant, & A. C. Simon (Eds.), *Fluency and disfluency across languages and language varieties* (pp. 9–14). Presses Universitaires de Louvain.

Derwing, T. M. (2017). L2 fluency development. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 246–259). Routledge.

Derwing, T. M., Munro, M. J., & Thomson, R. I. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics, 29*, 359–380.

Doe, T. (2021). Fluency development in an EFL setting: A one-semester study. *Language Teaching Research*. Advance online publication.

Dumont, A. (2018). *Fluency and disfluency: A corpus study of non-native and native speaker (dis)fluency profiles* (Unpublished doctoral dissertation). Université catholique de Louvain. http://hdl.handle.net/2078.1/198393

Dörnyei, Z. (1995). On the teachability of communication strategies. *TESOL Quarterly, 29*, 55–85.

Dörnyei, Z., & Kormos, J. (1998). Problem-solving mechanisms in L2 communication: A psycholinguistic perspective. *Studies in Second Language Acquisition, 20*, 349–385.

Finnish National Agency for Education (2019). *Lukion opetussuunnitelman perusteet* 2019 [National core curriculum for upper secondary schools 2019].

https://www.oph.fi/sites/default/files/documents/lukion_opetussuunnitelman_perusteet_2019.pdf

Guillot, M.-N. (1999). *Fluency and its teaching*. Multilingual Matters.

Götz, S. (2013). *Fluency in native and nonnative English speech*. John Benjamins.

Hanzawa, K. (2021). Development of second language speech fluency in foreign language classrooms: A longitudinal study. *Language Teaching Research* (online first).

Housen, A., Kuiken, F., & Vedder, I. (Eds.). (2012). *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*. John Benjamins.

Huensch, A., & Tracy-Ventura, N. (2017). L2 utterance fluency development before, during, and after residence abroad: A multidimensional investigation. *The Modern Language Journal, 101*, 275–293.

Kormos, J. (2006). *Speech production and second language acquisition*. Lawrence Erlbaum Associates.

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

Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. Routledge.

Lennes, M. (2002). *Total duration of labeled segment*s. *Praat script*. https://github.com/lennes/spect/blob/master/scripts/total_duration_of_labeled_segments.praat

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

Lintunen, P., & Eirola, V. (2023). Organizing an in-service teacher training module – societal interaction at its best. *EnglanTY blog*. https://blogit.utu.fi/englishturku/2023/06/05/organizing-an-in-service-teacher-training-module-societal-interaction-at-its-best/

Lintunen, P., & Mäkilähde, A. (2018). Short- and long-term effects of pronunciation teaching: EFL learners' views. In J. Volín & R. Skarnitzl (Eds.), *Pronunciation of English by Speakers of Other Languages* (pp. 46–72). Cambridge Scholars Publishing.

Lintunen, P., Mutta, M., & Peltonen, P. (2020). Synthesising approaches to second language fluency: Implications and future directions. In P. Lintunen, M. Mutta, & P. Peltonen (Eds.), *Fluency in L2 learning and use* (pp. 186–201). Multilingual Matters.

Mitrunen, M. (2023). Sallivuutta suulliseen ilmaisuun syntyperäisten puhetta havainnoimalla [More tolerance towards spoken expression by observing native speech]. In P. Lintunen, S. Mäkinen, & V. Eirola (Eds.), *Sujuvuutta suulliseen kielitaitoon: Näkökulmia kielten opetukseen* [Fluent oral proficiency: Perspectives on language teaching] (pp. 82–86). University of Turku.

Mora, J. C., & Valls-Ferrer, M. (2012). Oral fluency, accuracy, and complexity in formal instruction and study abroad learning contexts. *TESOL Quarterly, 46*, 610–641.

Peltonen, P. (2020). *Individual and interactional speech fluency in L2 English from a problem-solving perspective: A mixed-methods approach*. University of Turku.

Peltonen, P. (2021, July). *Development of individual and interactional L2 fluency*. Paper presented at the 30th EuroSLA conference 2021, University of Barcelona, Spain (online). https://eurosla30.dryfta.com/program-schedule

Peltonen, P., & Lehtilä, E. (2023). Sujuvuus oppimisen ja arvioinnin kohteena [Fluency as the target of learning and assessment]. In P. Lintunen, S. Mäkinen, & V. Eirola (Eds.), *Sujuvuutta suulliseen kielitaitoon: Näkökulmia kielten opetukseen* [Fluent oral proficiency: Perspectives on language teaching] (pp. 8–15). University of Turku.

Peltonen, P., & Lintunen, P. (2016). Integrating quantitative and qualitative approaches in L2 fluency analysis: A study of Finnish-speaking and Swedish-speaking learners of English at two school levels. *European Journal of Applied Linguistics, 4*, 209–238.

Peltonen, P., & Lintunen, P. (2019, August). *Short-term development of L2 fluency in formal instruction: A mixed methods study*. Paper presented at the 29th EuroSLA conference 2019, Lund University, Sweden. https://www.humlab.lu.se/fileadmin/user_upload/humlab/EuroSLA_Book_of_Abstracts.pdf

Rossiter, M. J., Derwing, T. M., Manimtim, L. G., & Thomson R. I. (2010). Oral fluency: The neglected component in the communicative language classroom. *The Canadian Modern Language Review, 66,* 583–606.

Segalowitz, N. (2010). *The cognitive bases of second language fluency*. Routledge.

Segalowitz, N., & Freed, B. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition, 26*, 173–199.

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

Tavakoli, P. (2018). L2 development in an intensive study abroad EAP context. *System, 72*, 62–74.

Tavakoli, P., Campbell, C., & McCormack, J. (2016). Development of speech fluency over a short period of time: Effects of pedagogic intervention. *TESOL Quarterly, 50*, 447–471.

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*, 330–349.

Tavakoli, P., Nakatsuhara, F., & Hunter, A.-M. (2020). Aspects of fluency across assessed levels of speaking proficiency. *The Modern Language Journal, 104*, 169–191.

Tavakoli P., & Wright, C. 2020. *Second language speech fluency: From research to practice*. Cambridge University Press.

Tonkyn, A. (2012). Measuring and perceiving changes in oral complexity, accuracy and fluency: Examining instructed learners' short-term gains. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 221–246). John Benjamins.

Towell, R. (2002). Relative degrees of fluency: A comparative case study of advanced learners of French. *International Review of Applied Linguistics in Language Teaching, 40*, 117–150.

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

Wolk, C., Götz, S., & Jäschke, K. (2020). Possibilities and drawbacks of using an online application for semi-automatic corpus analysis to investigate discourse markers and alternative fluency variables. *Corpus Pragmatics, 5*, 7–36.

Wood, D. (2006). Uses and functions of formulaic sequences in second language speech: An exploration of the foundations of fluency. *The Canadian Modern Language Review, 63*, 13–33.

**Appendix 1.** Example of a practice task

Listen to the learner's speech samples in English and in Finnish. Answer the following questions:

1. What's the learner's *speech rate* like in English/Finnish?
    - Think about the pace, rhythm, and stress patterns.
2. What are the learner's *pausing patterns* like in English/Finnish?
    - Think about the frequency, duration, and location of pauses.
3. What kinds of *resources* does the learner use to *maintain* fluency and avoid long silences in English/Finnish?
    - The resources can include filled pauses (*uh, um / öö*), drawls (sound lengthening), filler words (*like, you know, well, I mean… / niinku, tota, no*), and self-repetition.
4. Any other comments about the learner's fluency and individual speaking style:

Record yourself in English and in Finnish. Then listen to your own speech and answer the same questions* below.

(*questions 1–4, but from the student's perspective)

**Appendix 2a.** Fluency measures by participant (excluding FPs and repetitions reported in Figures 1 and 2) at T1

| Participant | Articulation rate | Number of SPs / min. | Mean length of SPs | Speech rate | Mean length of utterance |
|---|---|---|---|---|---|
| 1 | 203.65 | 31.43 | 0.78 | 144.32 | 6.35 |
| 2 | 189.13 | 46.56 | 0.83 | 114.85 | 4.00 |
| 3 | 211.17 | 51.54 | 1.19 | 104.37 | 4.00 |
| 4 | 260.47 | 45.02 | 0.72 | 169.26 | 5.59 |
| 5 | 221.42 | 41.47 | 0.81 | 142.05 | 5.25 |
| 6 | 230.26 | 37.97 | 0.82 | 151.54 | 5.94 |
| 7 | 235.83 | 29.48 | 1.09 | 153.66 | 7.68 |
| 8 | 227.16 | 44.26 | 0.86 | 139.29 | 5.04 |
| 9 | 213.95 | 27.86 | 0.76 | 157.89 | 7.41 |
| 10 | 168.69 | 27.21 | 0.77 | 124.86 | 5.96 |
| 11 | 203.09 | 33.85 | 0.93 | 133.20 | 5.77 |
| 12 | 193.75 | 49.05 | 0.88 | 112.69 | 3.85 |
| 13 | 236.48 | 49.60 | 0.79 | 143.17 | 4.66 |
| 14 | 233.73 | 29.11 | 0.71 | 173.81 | 7.80 |
| 15 | 184.54 | 23.30 | 0.72 | 144.43 | 7.62 |
| 16 | 241.67 | 46.25 | 0.81 | 148.49 | 5.10 |
| 17 | 182.23 | 31.57 | 0.63 | 136.72 | 5.61 |
| 18 | 209.63 | 31.57 | 0.80 | 147.34 | 6.38 |

Note: Participants 1–8 = control group, participants 9–18 = experimental group.

**Appendix 2b.** Fluency measures by participant (excluding FPs and repetitions reported in Figures 1 and 2) at T2

| Participant | Articulation rate | Number of SPs / min. | Mean length of SPs | Speech rate | Mean length of utterance |
|---|---|---|---|---|---|
| 1 | 215.83 | 21.48 | 0.86 | 164.88 | 9.57 |
| 2 | 190.98 | 38.40 | 0.83 | 124.72 | 4.85 |
| 3 | 207.50 | 53.86 | 0.85 | 117.90 | 3.74 |
| 4 | 238.60 | 37.67 | 0.70 | 165.96 | 6.11 |
| 5 | 220.62 | 35.63 | 0.65 | 159.46 | 5.96 |
| 6 | 258.18 | 42.15 | 0.86 | 161.15 | 5.94 |
| 7 | 229.54 | 19.65 | 0.82 | 180.75 | 11.17 |
| 8 | 200.79 | 42.54 | 0.57 | 142.76 | 4.54 |
| 9 | 221.07 | 36.85 | 1.08 | 132.80 | 5.77 |
| 10 | 177.93 | 35.59 | 0.70 | 125.96 | 4.89 |
| 11 | 190.89 | 39.36 | 0.78 | 126.01 | 4.73 |
| 12 | 188.27 | 41.30 | 0.99 | 112.18 | 4.45 |
| 13 | 223.97 | 41.51 | 0.83 | 142.09 | 5.29 |
| 14 | 218.62 | 26.17 | 0.82 | 161.24 | 8.09 |
| 15 | 186.36 | 24.65 | 0.85 | 138.15 | 7.38 |
| 16 | 255.70 | 37.91 | 0.81 | 169.26 | 6.59 |
| 17 | 188.48 | 34.41 | 0.72 | 133.50 | 5.25 |

| | | | | | |
|---|---|---|---|---|---|
| 18 | 197.30 | 40.91 | 0.90 | 122.29 | 4.72 |

Note: Participants 1–8 = control group, participants 9–18 = experimental group.

**Appendix 3a.** Frequencies of syllables, utterances, SPs, FPs, and repetitions by participant at T1

| Participant | Number of syllables | Number of utterances | Number of SPs | Number of FPs | Number of repetitions |
|---|---|---|---|---|---|
| 1 | 311 | 49 | 48 | 4 | 2 |
| 2 | 260 | 65 | 64 | 25 | 5 |
| 3 | 168 | 42 | 41 | 7 | 9 |
| 4 | 162 | 29 | 28 | 0 | 0 |
| 5 | 315 | 60 | 59 | 6 | 0 |
| 6 | 285 | 48 | 47 | 8 | 1 |
| 7 | 192 | 25 | 24 | 1 | 0 |
| 8 | 272 | 54 | 53 | 10 | 2 |
| 9 | 215 | 29 | 28 | 5 | 0 |
| 10 | 155 | 26 | 25 | 10 | 5 |
| 11 | 150 | 26 | 25 | 9 | 1 |
| 12 | 158 | 41 | 40 | 7 | 6 |
| 13 | 205 | 44 | 43 | 8 | 2 |
| 14 | 273 | 35 | 34 | 29 | 17 |
| 15 | 198 | 26 | 25 | 8 | 2 |

| | | | | | |
|---|---|---|---|---|---|
| 16 | 209 | 41 | 40 | 2 | 3 |
| 17 | 202 | 36 | 35 | 14 | 1 |
| 18 | 166 | 26 | 25 | 1 | 3 |

Note: Participants 1–8 = control group, participants 9–18 = experimental group.

**Appendix 3b.** Frequencies of syllables, utterances, SPs, FPs, and repetitions by participant at T2

| Participant | Number of syllables | Number of utterances | Number of SPs | Number of FPs | Number of repetitions |
|---|---|---|---|---|---|
| 1 | 201 | 21 | 20 | 5 | 0 |
| 2 | 189 | 39 | 38 | 16 | 4 |
| 3 | 131 | 35 | 34 | 7 | 2 |
| 4 | 171 | 28 | 27 | 1 | 0 |
| 5 | 161 | 27 | 26 | 4 | 1 |
| 6 | 196 | 33 | 32 | 0 | 2 |
| 7 | 257 | 23 | 22 | 3 | 0 |
| 8 | 118 | 26 | 25 | 4 | 0 |
| 9 | 150 | 26 | 25 | 3 | 0 |
| 10 | 225 | 46 | 45 | 15 | 9 |
| 11 | 194 | 41 | 40 | 3 | 3 |
| 12 | 196 | 44 | 43 | 10 | 8 |
| 13 | 259 | 49 | 48 | 2 | 0 |
| 14 | 259 | 32 | 31 | 2 | 5 |

| | | | | | |
|---|---|---|---|---|---|
| 15 | 310 | 42 | 41 | 13 | 9 |
| 16 | 290 | 44 | 43 | 4 | 9 |
| 17 | 126 | 24 | 23 | 6 | 2 |
| 18 | 217 | 46 | 45 | 0 | 0 |

Note: Participants 1–8 = control group, participants 9–18 = experimental group.

Address for correspondence

Pauliina Peltonen

University of Turku

Department of English

Arcanuminkuja 1, Building Arcanum, Office B341

20014 University of Turku

Finland