openjournals.nl
Dutch Journal of Applied Linguistics
https://www.dujal.nl | E-ISSN 2211-7253

The role of foreign language anxiety and task complexity on fluency in German learners of Dutch

Foekje Reitsma[1] and Esther Ruigendijk[1,2]
[1]*Department of Dutch Studies, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany* |
[2] *Cluster of Excellence "Hearing4all", Oldenburg, Germany*

Abstract This study investigated which aspects of fluency are related to foreign language anxiety and proficiency, and how this interacts with task complexity during a non-exam situation. Sixty-one low-intermediate German learners of Dutch completed a foreign language anxiety questionnaire, a proficiency test and two speech production tasks. Correlational analyses showed that anxiety was negatively related to number of mid-clause pauses in a complex task. Proficiency was positively related to numerous speed and breakdown fluency measures in a simple task. Mixed-effects models demonstrated that proficiency predicted two fluency measures. Task type positively predicted speed fluency. Anxiety was not a significant predictor of any fluency measure, which may be related to participants' relatively low anxiety level. This finding suggests that anxiety may not have a strong influence during speaking tasks that are not part of formal assessments.

Article history
Received: November 27, 2022
Accepted: August 8, 2024
Online: October 25, 2024

Corresponding author
Foekje Reitsma,
foekje.reitsma@uni-oldenburg.de

Author contributions
Funding information
Statement of interest
Statement of technology use
Supporting information
See p. 18

Keywords fluency, foreign language anxiety, task complexity, proficiency

1 Introduction

Highly proficient speakers produce language at a fast rate and with little effort. In contrast, learners of a second or foreign language (L2) often produce speech at a slower speech rate with more pauses and repairs. These disfluencies (i.e., pausing, slowing down or repairing) are occasionally assessed as a sign of lower proficiency in L2 research or in the classroom, but they may also be related to affective factors such as foreign language anxiety (FLA) (Tavakoli & Wright, 2020; Kormos, 2015). Horwitz et al. (1986) define FLA as "a distinct complex of self-perceptions, beliefs, feelings, and behaviors related to classroom language learning arising from the uniqueness of the language learning process" (p. 128). According to Horwitz et al. (1986), "anxious students feel a deep self-consciousness when asked to risk revealing themselves by speaking the foreign language in the presence of other people" (p. 129). Against this background, studies indicate that FLA negatively affects speech fluency (MacIntyre & Gardner, 1994; Pérez Castillejo, 2019, 2021; Sanaei et al., 2015; Bielak, 2022), by interfering with the cognitive processes that require attention control during speech processing (Eysenck et al. 2007).

How to cite Reitsma, F. and Ruigendijk, E. (2024). 'The role of foreign language anxiety and task complexity on fluency in German learners of Dutch', *Dutch Journal of Applied Linguistics* 13: DOI 10.51751/dujal13401

So far, especially L2 fluency in less cognitively demanding tasks has been investigated in relation to FLA (Hewitt & Stephenson, 2012; MacIntyre & Gardner, 1994; Pérez Castillejo, 2019, 2021; Phillips, 1992). Bielak (2022) investigated the link between FLA and fluency in a more demanding task among advanced L2 learners and found that proficiency was the strongest predictor of fluency, followed by much weaker predictive power of anxiety. To our knowledge, the relationship between FLA and fluency in low-intermediate L2 learners has not been investigated in relation to task complexity.[1] Lower participant proficiency (as compared to advanced L2 participants in Bielak, 2022) may cause FLA effects to be stronger due to less automatization (i.e., efficiency with which planning, selecting words and syntactic structures, and articulating message occur, Segalowitz, 2010) in L2 use. Studying the effect of task complexity is relevant for teachers who use these tasks in classroom practice and for researchers to better understand the role of FLA on fluency. The present study aims to fill this research gap by investigating which aspects of L2 oral fluency (i.e., speed, breakdown and repair fluency, more details will follow below) relate to FLA and how this interacts with task complexity.

2 Literature review

2.1 The speech production process and fluency

Speech fluency is a temporal feature of speech that can be defined both in a broad and a narrow sense (Lennon, 1990). In the broad sense, fluency refers to “spoken command of a foreign language” (Lennon, 1990, p. 389). In a narrow sense, it is used in research settings and L2 proficiency testing to refer to a specific aspect of oral proficiency that refers to the speed, smoothness and effortlessness of L2 speech (Chambers, 1997; Lennon, 1990). The present study concerns the narrow sense of fluency.

To understand how L2 fluency is achieved, it is necessary to look at the speech production process. Levelt (1989) or Levelt et al. (1999) distinguishes four main processes in his L1 speech production model: conceptualizing, formulating, articulating, and self-monitoring. In L1 (or highly proficient L2) speech, formulation (including lexical retrieval and grammatical encoding) and articulation are mainly automatic and allow for parallel processing with other modules (i.e., occurring virtually simultaneously and requiring little cognitive effort). Conceptualization (utterance planning) and monitoring processes require attention. These L1 speech production features explain “speed and fluidity, that is, a speaker’s fluency” (Tavakoli & Wright, 2020, p. 9).

For less proficient L2 speakers, not only conceptualization and monitoring, but also lexical retrieval, grammatical encoding and articulation might require conscious attention (de Bot, 1992; Kormos, 2006). The process will then occur less automatically, in a consecutive sequence (serial processing) and hence more slowly and effortfully. When L2 learners have access to a larger repertoire of lexical and grammatical knowledge and progress to a higher proficiency level, their lexical access and grammatical encoding becomes more automatic, and they therefore produce output of higher fluency. The fluency of L2 performance thus reflects the efficient functioning of speech production processes, namely conceptualization, formulation, articulation and monitoring (Kormos, 2006). This aspect of fluency is referred to as cognitive fluency (Segalowitz, 2010).

The degree of cognitive fluency is reflected in speech in the form of so called utterance fluency (Segalowitz, 2010), which can be further divided into speed fluency, or the number of words or (pruned[2]) syllables that are produced; breakdown fluency, or the number and length of pauses and repair fluency, which reflects hesitations and repairs (Lambert et al., 2017; Saito et al., 2018; Pérez Castillejo, 2019). Speed fluency reflects the degree of automatization of linguistic knowledge and the ease of accessing it (Pérez Castillejo, 2019, 2021; Bielak, 2022), which refer to the formulation/encoding (Levelt, 1989) aspect of cognitive fluency. Regarding breakdown fluency, pauses within clauses, which can reflect processing difficulties (Kahng, 2014), may be indicative of formulation /encoding difficulties/processes. Pauses between clauses may be indicative of content planning and conceptualization (Götz, 2013). Repair fluency is assumed to reflect the effectiveness of the monitoring system (Kormos, 1999).

In summary, while L1 speech processing is more automatized and stable, L2 speech processing, especially at lower proficiency levels, usually requires more control or conscious attention. Besides the degree of automatization, fluency can be affected by FLA, which will be discussed in the next section.

2.2 Foreign language anxiety, cognitive fluency and task complexity

Over time, attempts have been made to classify the FLA construct as trait anxiety (i.e., a permanent likelihood of experiencing anxiety regardless of the situation); state anxiety (which may arise in response to a situation at a particular time); or situation-specific anxiety (i.e., a personal predisposition to experience anxiety in particular situations, Baran-Lucarz, 2022). Since a pioneering article by Horwitz et al. (1986), FLA has been treated as a situation-specific type of anxiety (MacIntyre & Gardner, 1989). Recently, Horwitz (2017) characterized a person experiencing FLA as "having the *trait* of feeling *state* anxiety when participating in (or sometimes even thinking about) language learning and/or use" (p. 33). Dewaele (2007, pp. 405–406) explains that FLA is "probably situated halfway between trait, situation-specific anxiety and state, more sensitive to environmental factors than personality traits and yet more stable than states since it remains relatively stable across languages."

One explanation for the idea that FLA may influence the speech production process (i.e., more specifically fluency) comes from the Processing Efficiency Theory by Eysenck and Calvo (1992). They suggest that anxiety impairs the efficiency of cognitive processing during L2 speech production by posing additional demands on working memory resources and attentional processes.

The current study investigates the relationship between FLA and utterance fluency. Studies examining this link are limited. Sanaei et al. (2015) found strong negative correlations between FLA and speed fluency (i.e. more anxious participants had shorter runs between pauses and shorter phonation time) during a non-exam situation. Pérez Castillejo's (2019) examined how FLA and proficiency relate to L2 fluency during a final oral exam in learners with low-intermediate proficiency level. The results indicated medium and strong negative relationships between FLA and those fluency measures that reflect speech formulation/encoding (i.e. more anxious participants had a smaller phonation time, shorter runs between pauses and paused more frequently within clauses). The negative relationship between FLA and the length of pauses between clauses, which is mainly related to message conceptualization, was weaker. Regression analyses showed that FLA was a stronger predictor of fluency than proficiency. Whereas Pérez Castillejo (2021) replicated Pérez Castillejo's (2019) results regarding the relationship between FLA and the fluency of speech formulation/encoding (i.e., mean length of run and number of mid-clause pauses), their results regarding the relationship between FLA and conceptualization fluency (the length of pauses) were not replicated, indicating that there is more evidence that formulation/encoding is affected by anxiety than conceptualization is. Pérez Castillejo (2021) also found that repair fluency was very weakly related to FLA and proficiency, probably due to the low proficiency level of participants. Proficiency was however a much stronger predictor of L2 fluency than FLA when fluency was measured in a task following another similar task, that is, in a condition of prior L2 processing (understood as L2 use earlier in discourse).

In addition to anxiety, cognitive fluency may also be influenced by task complexity. According to Robinson's Cognition Hypothesis (2001), for instance, cognitively demanding tasks focus L2 speakers' attention on language form, to increase the grammatical accuracy and linguistic complexity of their L2 production at the cost of fluency (Michel et al., 2007; Robinson, 2001). The Limited Attentional Capacity Model (Skehan & Foster, 2001) on the other hand suggests that complex tasks may negatively affect both grammatical accuracy and fluency. Skehan and Foster (2001) argue that learners focus more on meaning than on form, and more so in cognitively demanding tasks. Both Robinson (2001) and Skehan and Foster (2001) thus predict that task complexity would negatively impact fluency. De Jong et al. (2012) examined how task complexity affected L2 speaking performance. Results showed that complex tasks led to more pausing and repair behavior than simple tasks, but no significant difference in speed fluency was found. Regarding lexical complexity, L2 speakers produced a wider range of words in complex tasks compared to simple tasks.

The effect of task complexity may vary depending on the proficiency and anxiety level of the speaker. Since the automaticity of linguistic encoding is related to proficiency (Kormos, 2006; Segalowitz, 2010), more proficient L2 speakers may have more attentional resources available to conceptualize the message than less-proficient speakers, who must divide their attention between conceptualization, formulation and monitoring (Lambert et al., 2017). Additionally, the effect of task complexity on speech processing might vary depending on the anxiety level of the speaker completing the task. According to Eysenck and Calvo (1992), the more cognitively complex the task, the stronger the effect of anxiety on task performance. Robinson's (2001) Cognition Hypothesis also states that individual learner differences, such as anxiety, have a greater influence on spoken task performance if tasks are more cognitively complex. To date, only Bielak (2022) examined the relationship between FLA and L2 fluency in advanced L2 learners. Participants completed both a group task and a monologue (i.e., complex task). Numerous negative correlations between FLA and fluency measures, reflecting formulation (i.e., more anxious participants had shorter runs and lower articulation rate), and conceptualization (more end-pauses), were found in the group task, and only one negative correlation (more end-pauses) was found during the complex task. Regression analyses showed that proficiency was a stronger predictor of fluency than FLA. This could be attributed to the advanced proficiency level of the participants and consequently more automatized L2 use. A limitation of this study was that FLA was measured twice, after the group and complex tasks, but fluency was measured only once, during the complex task; therefore, it is not certain whether the correlations between the first anxiety measurement and fluency are reliable.

Based on these previous studies (Sanaei et al., 2015; Pérez Castillejo's, 2019, 2021) one might expect that anxiety affects attentional processes, which in turn affects fluency in L2. Furthermore, based on Robinson (2001), Eysenck and Calvo (1992) and Bielak (2022) it is expected that anxiety will negatively affect speech fluency during a cognitively demanding task.

The present study investigates the effect of anxiety on speed, breakdown and repair fluency in two types of tasks, one simple and one complex, among low-intermediate L2 learners. On the basis of previous research, we hypothesize that a cognitively more demanding task negatively affects speed fluency of anxious learners. With respect to breakdown fluency, we expect that complex tasks elicit more pausing behavior. Considering repair fluency, we hypothesize that a cognitively demanding task will lead to fewer self-corrections and reformulations by anxious learners. The idea underlying this hypothesis is that at lower proficiency levels, when L2 knowledge is not automatized, learners' attentional capacity can be strained too much by other processes, particularly formulation (Kormos, 2000; Pérez Castillejo, 2021), so that there is no room left to make corrections.

Our research question is therefore as follows: "To what extent is there a relationship between speed, breakdown and repair fluency measures and FLA and proficiency, and how does this interact with task complexity?".

3 Method

In order to answer our research question, we performed two speech production tasks with a group of German learners of Dutch to examine their fluency and related this to their level of anxiety. Participants, materials, as well as a detailed description of the analyses are presented in the following section.

3.1 Participants

The participants were 61 German learners of Dutch (20 male and 41 female) from three secondary schools, all of whom were in their fifth year of Dutch. The L2 learners reported German as their L1. Six participants had two L1s: three German/*Niederdeutsch*, one German/Russian, one German/Kurdish and one German/Ukrainian. Their ages ranged from 14 to 16 with an average age of 15.2 years. All students participated in this study voluntarily with signed informed consent.

3.2 Anxiety measure

The original FLCAS questionnaire (Horwitz et al., 1986) consists of 33 items with a 5-point Likert scale concerning learners' feelings and behavior in the foreign language classroom, and in conversation with native speakers. This instrument has been validated and widely used in language learning research.

Mak (2011) found that the FLCAS tests five subconstructs of anxiety: (a) speech anxiety and fear of negative evaluation, (b) uncomfortableness when speaking with native speakers, (c) negative attitudes toward the L2 class, (d) negative self-evaluation, and (e) fear of failing the class or the consequences of personal failure. Considering our research question only three of these subconstructs with in total 21 items were of relevance and used in our study: the first two (a) and (b), as well as (d). The "negative attitudes toward the Dutch class" subconstruct was not selected because the questionnaire was completed in the first weeks of the school year and hence, participants' responses were less likely to be influenced by their attitudes toward the class. The "fear of failing the class or the consequences of personal failure" subconstruct also was not relevant because the speech production tasks did not constitute an exam and not related to the class in that sense.

A native German speaker with a Master's degree in English translated the original English version into German. A professional translator subsequently back-translated the questionnaire and confirmed that the English and German versions were completely comparable. The questionnaire was then piloted with participants not involved in the study to test whether the items were clearly formulated. Feedback showed that one item had a translation error which was subsequently corrected. Participants completed the questionnaire in approximately five minutes. The reliability of the questionnaire as measured by Cronbach's alpha was 0.89, which indicates a high level of internal consistency.

3.3 Proficiency measure

The participants completed the Dutch version of the LexTALE (Lemhöfer & Broersma, 2012), a five-minute, 60-item vocabulary test, in which they must indicate whether or not a string of letters represents an existing word in the language. The test measured participants' general proficiency in Dutch. A benefit of the LexTALE is that the ranges of scores are associated with the proficiency levels as described in the Common European Framework of Reference (CEFR) (Council of Europe, 2001; see Lemhöfer and Broersma (2012) for the correspondence between LexTALE and CEFR).[3] According to Lemhöfer and Broersma (2012), LexTALE scores of 80–100 % represent CEFR level C1–C2, 60–80 % level B2, and below 59 % level B1 and lower. Cronbach's alpha assessed the LexTALE's reliability and was 0.71, which indicates an acceptable level of internal consistency.

3.4 Speech production tasks

Speech data was collected using two tasks that differed in complexity. To operationalize complexity, we followed De Jong et al. (2012). In their opinion, complex tasks contain more elements than simple tasks; complex tasks concern a more general topic as opposed to simple tasks, which concern topics of personal life; and complex tasks involve more abstract notions as opposed to simple tasks, which involve mostly concrete notions. In our simple task, the participants discussed their weekend activities. This is a concrete topic of personal life, and the task consists of one element: "What did you do last weekend?". In our complex task, participants were asked to persuade the class of their opinions and arguments about smartphones in the classroom. This is a more general topic; it is more abstract and consists of two elements (providing their opinion and arguments; see Appendix A). The regular class teachers confirmed that the tasks were appropriate for the participants' Dutch level. Moreover, 14 German students of Dutch piloted the tasks. Their feedback showed that the instructions were clear and that the more cognitively demanding task was perceived as more complex than the simple task.

3.5 Procedure

The participants completed the FLCAS questionnaire, as well as the demographic and language background questionnaire and the proficiency test some weeks before the speech production tasks. We measured anxiety independently of the speaking tasks, to be better able to compare our data to similar studies (Sanaei et al., 2015; Pérez Castillejo, 2019).

The data from the oral tasks were collected on regular school days where the participants met individually with the researcher whom they had already met during the first session (completion of questionnaires and proficiency test). The participants were given the simple and complex tasks on paper and, after checking for comprehension, were instructed to answer without time to prepare. They were asked to speak for two minutes per task and told that their performances were audio recorded. In both tasks, participants spoke only to the researcher, not to a classmate. The role of the researcher during the performance was limited to nodding to show understanding. The tasks had not been practiced before. For the sake of comparability, the tasks were always taken in the same sequence; first the simple task and then the complex task. All speech samples were recorded on an Olympus LS 10 digital voice recorder.

3.6 Data coding and analysis

In the simple task, participants spoke on average for 58.00 seconds (SD = 25.41) from the first to the last syllable, and in the complex task for 54.33 seconds (SD = 22.39). All speech recordings of the two tasks of the 59 participants (roughly two hours of speech in total) were transcribed in detail, including information regarding pauses and repairs. Two of the 61 participants were absent when the speech data was collected. Three datasets were discarded due to technical problems. One recording lasted five seconds: we excluded it from the analyses, as fluency measures are unstable for short speech samples. This resulted in 56 datasets from the simple task and 58 from the complex task. These 114 recordings were transcribed into Analysis of Speech (AS) units following the procedures of Foster et al. (2000). The AS unit is defined as “a single speaker’s utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clause(s) associated with either” (Foster et al., 2000, p. 365).

As a first step, the “textgrid-to-silences” function in PRAAT (Boersma & Weenink, 2015) was used to automatically mark silent pauses. Since the background noise was occasionally disturbing automatic detection of silent pauses, the identified pauses were manually checked and adjusted as necessary. Only silent pauses (i.e., greater than or equal to 0.25 seconds) and filled pauses were marked as pauses. Short “micropauses” (Riggenbach, 1991) that are irrelevant for measures of L2 fluency were excluded (De Jong & Bosker, 2013). All filled pauses were manually marked.

For each pause it was determined whether it was mid-clause or end-clause through careful listening to the recordings and examination of the transcript, which had been marked with clause boundaries. Mid-clause pauses were filled and silent pauses occurring in the middle of clauses. Pauses between clauses and when a conjunction occurred between clauses were considered to be end-clause pauses.

For each speech recording, syllables were counted manually. Syllables were counted for the purpose of calculating speech rate, articulation rate and mean length of run (i.e., stretches of speech uninterrupted by pauses > 0.25 seconds). The number of self-corrections and reformulations during each performance was also counted manually.

Regarding fluency measures, it has been argued (Bosker et al., 2013; Huensch & Tracy-Ventura, 2016; Hunter, 2017; Skehan, 2015, Tavakoli et al., 2020) that composite and pure fluency measures need to be distinguished. Speech rate, mean length of run and phonation time ratio are described as composite measures because they combine two or more of the aspects of fluency. A measure such as mean length of run combines speed and breakdown fluency. In contrast, pure measures relate to one of the three aspects of fluency: speed, breakdown, or repair. For example, articulation rate relates to speed fluency. Pure measures have been argued to tell more about the underlying processes involved in speech production (Huensch and Tracy-Ventura, 2017; i.e., linking specific fluency measures to stages of speech production; Lambert et al., 2017).

The present study used several different (composite and pure) fluency measures (see Table 1). Sanaei et al. (2015), Pérez Castillejo (2019, 2021) and Bielak (2022) included a limited number of fluency measures. To capture a more complete picture of how anxiety relates to speed, breakdown and repair phenomena in a simple and complex task, we used a broad range of fluency measures (see Table 1 for details on their operationalization). For comparison purposes fluency measures from FLA and fluency studies were used, namely: speech rate, articulation rate; phonation time ratio; mean length of run (Sanaei et al., 2015); number of mid-clause and end-clause pauses; mean length of mid-clause; mean length of end-clause pauses (Pérez Castillejo, 2019; Bielak, 2022) and number of reformulations and self-corrections (Pérez Castillejo, 2021). In addition, we used other commonly used speed, breakdown and repair fluency measures (i.e., number of silent and filled pauses per minute, number of pauses per minute, and mean length of pauses) based on earlier fluency studies (See Kahng, 2022).

Data for six participants (i.e., 12 recordings; 10.53 % of the data) was re-examined for syllable count by an independently trained teacher/researcher. Additionally, the rater analyzed and coded 10 % of the data again for pause type, pause length, pause location, self-corrections and reformulations. Interrater reliability was 98 % agreement on syllable count, 98 % on pause type, 90 % on pause length, 92 % on pause location and 92 % on self-corrections and reformulations, demonstrating high interrater reliability.

3.7 Statistical analysis

As a first step, we checked the data (fluency measures, anxiety, and proficiency) for normality. The results of Shapiro-Wilk test indicated that the fluency measures: mean length of run, number of filled pauses, mean length of pauses, mean length of mid and end-clause pauses, and reformulations were not normally distributed ($p < 0.05$).

To explore the research question, a correlation analysis was run with the L2 learners' data on speed, pause phenomena, repair, anxiety and proficiency scores. Due to non-normality of some fluency measures, we decided to use Spearman rank-order correlation test. Based on this exploration, we decided to run regression analyses. Prior to the regression analyses, the assumptions of LMM were checked, including linearity, absence of collinearity, homoscedasticity, and normality and no violations were noted. Then a mixed-effects analysis was conducted, using the lme4 package (Bates et al., 2015) in R

Table 1 Fluency measures and their operationalization (Sample time refers to total duration of the recording)

Measure	Operationalization
Speed fluency	
Speech rate	Total number of syllables divided by total sample time and multiplied by 60.
Mean length of run	Total number of syllables divided by total number of runs.
Articulation rate	Total number of syllables divided by total speaking time (excluding all pauses > 0,25 sec) and multiplied by 60.
Phonation time ratio	Total speaking time divided by total sample time and multiplied by 100.
Breakdown fluency	
Number of silent pauses	Total number of silent pauses divided by total speaking time and multiplied by 60.
Number of filled pauses	Total number of filled pauses divided by total speaking time and multiplied by 60.
Number of pauses	Total number of pauses divided by total speaking time and multiplied by 60.
Number of mid-clause pauses	Total number of mid-clause pauses divided by total sample time and multiplied by 60.
Number of end-clause pauses	Total number of end-clause pauses divided by total sample time and multiplied by 60.
Mean length of pauses	Total length of pause time divided by number of pauses.
Mean length of mid-clause pauses	Total length of mid-clause pauses divided by number of mid-clause pauses.
Mean length of end-clause pauses	Total length of end-clause pauses divided by number of end-clause pauses.
Repair fluency	
Number of reformulations and self-corrections	Total number of reformulations and self-corrections divided by total sample time and multiplied by 60.

Version 4.0.2 (R Core Team, 2020). First, we tested interactions, for every fluency measure that correlated with anxiety and/or proficiency (i.e. speech rate, number of mid-clause pauses, number of pauses, number of silent pauses and phonation time ratio) and center the predictors involved in interactions (i.e., Anxiety, Proficiency and Task type). However, for all five fluency measures the models turned out to be non-significant and did not improve the model fit. We also added random slopes with interactions, but we received a warning that the data was not large enough to estimate random effects. Finally, for each fluency variable, we built a model with Anxiety, Proficiency and Task type as fixed effects and Subject as a random effect and then added School as a fixed effect. The fixed-effect factor, Task type, was added to control for differences between performing the simple

and complex tasks. School was added as a fixed effect to control for possible differences between the three schools that participated in the study. The variable Subject was added as a random effect to control for differences between participants. The models were subsequently compared; in essence, we tested whether to include the fixed effect School or not, and if the first model was preferable compared to second one. Model comparisons were assessed using the ANOVA (Model 1, Model 2) function in R.

4 Results

In this section, the results are summarized. Participants' FLCAS scores ranged from 23 to 72 or (when dividing the number of questions by the score) from 1.10 to 3.43. According to Horwitz (2013), averages below 3 indicate a low anxiety level, averages of approximately 3 a moderate level of anxiety and averages in the range near 4 to 5 as fairly anxious. The participants' mean score (2.23) can thus be considered to be at a low anxiety level. The scores on the proficiency test (LexTALE) showed that participants' scores ranged from 38.75 % to 72.50 %, with a mean score of 55.50 % (SD = 6.06). This indicates that their Dutch proficiency level was situated between B1 (or lower) and B2 on the CEFR.

Table 2 contains descriptive statistics for all fluency measures in the simple and complex speech production task. The results indicate that the participants spoke faster in the complex (mean = 131.28; SD = 27.06) than in the simple task (mean = 123.85; SD = 22.33). They articulated on average 204.67 (simple task) versus 211.23 (complex task) syllables, spoke for approximately 63 % (simple task) and 65 % (complex task) of the total performance time (i.e., phonation time ratio) and paused after uttering three to ten syllables (i.e., mean length of run). In both tasks, participants used more silent than filled pauses. Pauses were more frequent within than between clauses, but mid-clause pauses were shorter than end-clause pauses (simple task: 0.61 versus 0.79; complex task: 0.62 versus 0.72). On average, participants made more reformulations and self-corrections per minute in the simple task than in the complex task (simple task: mean = 1.72, SD = 1.82; complex task: mean = 1.67, SD = 1.63). Overall, the complex task resulted in faster speech and articulation rate, longer runs, more phonation time, more and longer mid-clause pauses, and less and shorter end-clause pauses than the simple task.

Table 3 displays correlation coefficients between anxiety and the fluency measures, which basically show low values. We found only one significant correlation between anxiety and fluency: number of mid-clause pauses in the complex task ($r = .317$, $p < 0.05$). The negative correlation between anxiety and proficiency (LexTALE scores; $r = -.128$, $p = 0.32$) was not statistically significant. Higher proficiency (LexTALE scores) was significantly correlated in the simple task with higher speech rate (simple task: $r = .325$, $p < 0.05$); higher phonation-time ratio (PTR; $r = .329$, $p < 0.05$); fewer silent pauses ($r = -.266$, $p < 0.05$), fewer pauses ($r = -.320$, $p < 0.05$); and fewer mid-clause pauses ($r = -.292$, $p < 0.05$).

Table 2 Descriptive statistics (M = mean; SD= Standard Deviation, Median, Min-Max = range) for the fluency measures in the simple and complex task

	Simple task (*N*=56)		Complex task (*N*=58)	
	M (SD) Median	**Min–Max**	**M (SD) Median**	**Min–Max**
Measures				
Speed fluency				
Speech rate	123.85 (22.33) 123.82	77.08–168.50	131.28 (27.06) 131.83	48.37–184.83
Mean length of run	4.89 (1.55) 4.55	2.95–9.31	5.00 (1.33) 4.74	3.05–9.88
Articulation rate	204.67 (22.98) 206.77	151.53–249.16	211.23 (22.48) 212.07	159.58–254.32
Breakdown fluency				
Phonation time ratio	63.01 (9.04) 62.53	46.07–81.50	64.85 (9.25) 66.49	25.83–82.11
Number of silent pauses	40.86 (12.97) 40.26	3.32–68.32	39.40 (9.92) 38.21	16.56–56.86
Number of filled pauses	12.57 (9.44) 10.56	0.00–47.92	12.60 (9.00) 11.43	0.00–35.33
Number of pauses	53.43 (17.80) 54.21	22.56–101.60	52.00 (13.30) 52.44	21.39–82.71
Number of mid-clause pauses	17.57 (6.05) 15.71	4.77–32.06	18.76 (5.89) 18.66	6.79–30.91
Number of end-clause pauses	14.74 (4.25) 14.04	6.84–29.89	14.08 (4.08) 14.36	5.86–24.42
Mean length of pauses	0.69 (0.13) 0.67	0.48–1.10	0.67 (0.26) 0.62	0.37–2.15

Table 2 Descriptive statistics for the fluency measures in the simple and complex task (*cont.*)

	Simple task (*N*=56)		**Complex task (*N*=58)**	
	M (SD) Median	**Min–Max**	**M (SD) Median**	**Min–Max**
Mean length of mid-clause pauses	0.61 (0.15) 0.59	0.37–1.04	0.62 (0.25) 0.55	0.28–1.72
Mean length of end-clause pauses	0.79 (0.22) 0.71	0.47–1.80	0.72 (0.31) 0.66	0.41–2.36
Repair fluency				
Number of reformulations and corrections	1.72 (1.82) 1.35	0.00–6.58	1.67 (1.63) 1.00	0.00–4.68

Table 3 Correlations between anxiety (FLCAS), proficiency (LexTALE) and fluency measures

	Anxiety		**Proficiency**	
	Simple task	**Complex task**	**Simple task**	**Complex task**
Measures				
Speed fluency				
Speech rate	–.111	–.069	**.325****	.092
Mean length of run	–.110	–.122	.250	.072
Articulation rate	–.060	–.035	.015	–.002
Breakdown fluency				
Phonation time ratio	–.124	–.040	**.329****	.087
Number of silent pauses	.137	.082	**–.266****	–.084
Number of filled pauses	.124	.123	–.237	–.175
Number of pauses	.166	.144	**–.320****	–.181
Number of mid-clause pauses	.117	**.317****	**–.292****	–.253
Number of end-clause pauses	.159	–.167	–.045	.056
Mean length of pauses	–.074	–.086	–.072	.083
Mean length of mid-clause pauses	–.021	.024	–.132	.039
Mean length end-clause pauses	–.079	–.130	–.012	.127

Table 3 Correlations between anxiety and fluency measures (*cont.*)

	Anxiety		**Proficiency**	
	Simple task	**Complex task**	**Simple task**	**Complex task**
Repair fluency				
Number of reformulations and self-corrections	.024	.139	–.164	–.103

*p < 0.01, **p < 0.05

Table 4 Results of mixed-effects models on factors affecting speech rate

Fluency measure	**Fixed effects: Factor**	***Estimate***	***SE***	***p***	**Random effects: Factor**	**Variance**
Speech rate	Intercept	92.61	30.69	0.001**	Participant	344.5
	Proficiency	0.72	0.48	0.137		
	Anxiety	–0.16	0.26	0.537		
	Task type	6.26	3.11	0.05*		

Plonsky and Oswald (2014) recommend the following benchmarks for the interpretation of effect size in correlation coefficients: close to 0.25 small, 0.40 medium, and 0.60 large. The effect size of the statistically significant correlations (r = –.266 to r =.329) in the current study are thus considered small.

Regression analyses were run to further test the relative contributions of the factors Anxiety, Proficiency and Task type to the variance in the five fluency measures that (at least weakly) correlated with anxiety or proficiency (i.e., speech rate, phonation time ratio, number of silent pauses, number of pauses and mid-clause pauses).

Tables 4, 5 and 6 present a summary of the models on factors affecting speech rate, number of mid-clause pauses and number of pauses that gave a significant effect (for a summary of the models and model comparisons for which we did not find a statistical effect, see Appendices B and C).

Table 4 presents the resulting model on factors affecting speech rate. Adding the fixed factor School did not significantly improve the model ($\chi^2(2) = 2.33$, $p = 0.31$). Task type emerged as the strongest predictor of speech rate, with the largest coefficient. The more complex the task, the higher the speech rate. Proficiency and Anxiety did not significantly predict speech rate.

Table 5 shows the resulting model on factors affecting number of mid-clause pauses. The results from the model comparison indicate that adding of the factor School does

Table 5 Results of mixed-effects models on factors affecting number of mid-clause pauses

Fluency measure	Fixed effects: Factor	*Estimate*	*SE*	*p*	Random effects: Factor	Variance
Number of mid-clause pauses	Intercept	25.15	6.89	<0.001	Participant	15.20
	Proficiency	−0.23	0.11	0.038		
	Anxiety	0.10	0.06	0.084		
	Task type	1.41	0.79	0.081		

Table 6 Results of mixed-effects models on factors affecting number of pauses

Fluency measure	Fixed effects: Factor	*Estimate*	*SE*	*p*	Random effects: Factor	Variance
Number of pauses	Intercept	76.45	18.58	<0.001	Participant	117.6
	Proficiency	−0.58	0.29	0.048		
	Anxiety	0.18	0.15	0.242		
	Task type	−0.82	2.03	0.689		

not improve the model ($\chi^2(2) = 1.71$, $p = 0.42$). Proficiency was significantly and negatively related to number of mid-clause pauses. The more proficient the speaker, the less mid-clauses pauses they use. Anxiety and Task type were not found to be significant predictors.

Table 6 presents the results of mixed-effects models on factors affecting number of pauses. The results from the model comparison show that adding the fixed factor School does not significantly improve the model ($\chi^2(2) = 0.44$, $p = 0.80$). Proficiency was negatively linked to number of pauses; the more proficient the speaker, the less pauses the speaker used.

To sum up, modeling demonstrated that Task type was a significant predictor of speech rate. Proficiency emerged as predictor of number of mid-clause pauses and number of pauses. Anxiety was not a significant predictor of any fluency measures. Furthermore, in all model comparisons, the addition of the fixed-effect School did not significantly improve the models.

5 Discussion

This study set out to answer the research question to what extent there is a relationship between speed, breakdown, and repair fluency measures, FLA, and proficiency, and how this interacts with task complexity. Correlational analyses showed that FLA was negatively related to one breakdown fluency measure, while proficiency was positively related to numerous speed and breakdown fluency measures. Regression analyses indicated that proficiency was a stronger predictor than FLA.

5.1 The relationship between FLA, proficiency and L2 fluency

The correlation analysis results that served as a first exploration of the data showed that more anxious speakers used significantly more mid-clause pauses in the complex task, but with a small effect size. The results generally support the theoretical claims of the Cognition Hypothesis (Robinson, 2001) and Processing Efficiency Theory (Eysenck & Calvo, 1992) regarding the negative role of anxiety in L2 speech processing, as well as previous research (Pérez Castillejo, 2019, 2021). In contrast, Bielak (2022) found no significant correlations between FLA and mid-clause pauses in a cognitively demanding task among advanced L2 learners. Mid-clause pauses reflect cognitive efficiency during formulation. Especially lexical retrieval and syntactic encoding (Kahng, 2014) may require attention. The participants in the present study had generally lower proficiency, which suggests that their ability to formulate may require greater attention control (Kormos, 2006). Thus, any interference that limits attention control during L2 speech production, such as FLA, may impact pausing behavior within clauses (Pérez Castillejo, 2021).

There were positive correlations between proficiency and fluency (i.e. speech rate and phonation time ratio) and negative correlations (i.e. number of silent and mid-clause pauses and number of pauses) in the simple task. These speed and pausing phenomena show how effectively L2 knowledge can be accessed and are linked to message formulation and encoding (Kahng, 2014). In contrast, fluency measures that may be associated with message conceptualization such as number of end-pauses were not related to proficiency in a statistically significant way. For L2 speakers, especially at lower proficiency levels, formulation and encoding may require more attention control than conceptualization (De Bot, 1992; Kormos, 2006; Segalowitz, 2010). This could explain why proficiency relates differently to the fluency measures associated with each of these processes. Note that the task we used, LexTALE has been validated for English, but not yet for Dutch, so especially with respect to assigning scores to the European Reference Framework, we cannot form any strong conclusions. The scores on the task can be used to correlate individual variation in proficiency though, which we have done here.

Furthermore, repair fluency was not significantly related to FLA or proficiency. This is in line with the findings of Pérez Castillejo (2021). Self-repairs may reflect that attentional resources are being used for monitoring (Kormos, 2000), but at lower proficiency levels, when L2 knowledge is not automatized, attention may be needed for other processes (particularly formulation), and hence FLA or proficiency may not significantly affect monitoring.

5.2 Predictors of L2 fluency

We additionally investigated the contributions of the predictors (i.e., FLA, proficiency and task type) on the fluency measures. First, proficiency was a significant predictor of two fluency measures, whereas FLA was not a significant predictor of any fluency measure. In Pérez Castillejo's study (2019) and its replication (2021) on the relationship between FLA, proficiency and fluency during a final oral exam, the roles seem reversed: FLA was a significant predictor of the fluency measures analyzed and proficiency was not. Explanations for the contradictory results could be that in the present study, the tasks were not performed during an exam situation, and the participants' anxiety levels were lower compared to the levels measured in previous research (Bielak, 2022; Pérez Castillejo, 2019, 2021; Saito et al., 2018).

Recall that the complex task resulted in a faster speech rate and articulation rate, longer runs, more phonation time, more and longer mid-clause pauses, and fewer and shorter end-clause pauses than the simple task. We found a significant positive effect of the predictor task type on speech rate. The more complex the task, the higher the speech rate. It is somewhat surprising that a cognitively demanding task results in a higher speech rate, as the Cognition Hypothesis (Robinson, 2001) and Limited Attentional Capacity Model (Skehan & Foster, 2001) predict that these tasks may negatively affect fluency. One possible explanation is that the complex task may not have been sufficiently complex, thus not resulting in more formulation problems. Another explanation is that participants had already used the L2 in the simple task and benefited from it while performing the complex task. According to McDonough and Trofimovich (2008), earlier activation of language material facilitates later access to and use of related linguistic knowledge, which can lead to improved performance. This is also what Pérez Castillejo (2021) and Bielak (2022) found in their research.

While this study is one of the first to address a broad range of fluency measures and address the influence of FLA, proficiency and task complexity, it also has some limitations. First, the participants produced relatively short speaking samples. The participants were instructed to speak for two minutes for each task, but on average they did not speak for more than one minute. Another limitation is that the tasks were always taken in the same sequence. This should have been varied to avoid the effect of novelty (better performance) in the first task and maybe the effect of repetitiveness (worse performance) in the second task. Future research might compare whether it makes a difference to vary the sequence of the simple and complex tasks. Finally, in the absence of L1 data from the participants, it is difficult to claim that the results obtained here are only due to the speakers' L2 fluency behavior and not affected by individual differences in their L1 speaking style (De Jong et al., 2015).

As far as pedagogical implications are concerned, the results of this study suggest that FLA may not have a strong influence during speaking tasks that are not part of formal assessments. This could encourage teachers to base grades for speaking proficiency on classroom observations as well, rather than only on formal tests (Bielak, 2022).

6 Conclusion

In conclusion, the current study examined the extent to which a relationship exists between speed, breakdown and repair fluency measures and FLA and proficiency, as well as how this interacts with task type. Based on previous research (e.g., Eysenck & Calvo, 1992; Pérez Castillejo, 2019, 2021; Robinson's Cognition Hypothesis, 2001; Bielak, 2022), we expected a cognitively demanding task to negatively relate to speed fluency and require more pausing behavior and fewer self-corrections and reformulations by anxious speakers. The expectations about the relationship between FLA and fluency were partially confirmed for breakdown fluency. The results indicated that FLA was a negative but not a significant predictor of number of mid-clause pauses, which may be related to participants' relatively low anxiety level. In contrast, proficiency related to numerous speed and breakdown fluency measures and was found to be a significant predictor of two breakdown fluency measures: number of mid-clause pauses and number of pauses. Unexpectedly, task type positively predicted speed fluency, suggesting that a cognitively demanding task led to more rather than less fluent performance. This finding suggests that the earlier activation of the L2 during the simple task could facilitate later access and use of L2 during the complex task.

Author contributions

Foekje Reitsma, conceptualization, methodology, investigation, data analysis, writing – original draft, writing – review and editing; Esther Ruigendijk, supervision, conceptualization, methodology, writing – review and editing.

Funding information

No funding was received for this study.

Statement of interest

The authors declare no conflicting interests.

Statement of technology use

No AI-based generative technology was used in the preparation of this manuscript and the execution of the research that the manuscript reports upon.

Supporting information

Appendix A – Speech production tasks. Appendix B – Results of mixed-effects models on factors affecting number of silent pauses and phonation time ratio. Appendix C – Model comparisons.

Notes

1 But note that Robinson (2007) focus on the interaction between input, processing and output anxiety (MacIntyre & Gardner, 1994) and language production (use of complex speech structures).

2 Pruned syllables means syllables excluding those that are repeated, reformulated or replaced.

3 Lextale is a test for advanced learners, it can distinguish between C1–C2, B2 and B1 or lower and is validated only for English so far, and so are the corresponding CEF levels. We therefore use the scores with care, since it may not be the same for Dutch.

References

Baran-Lucarz, M. (2022). Language Anxiety. In T.M. Derwing (Ed.). The *Routledge handbook of second language acquisition and speaking* (pp. 83–96). Routledge.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.

Bielak, J. (2022). To what extent are foreign language anxiety and foreign language enjoyment related to L2 fluency? An investigation of task-specific emotions and breakdown and speed fluency in an oral task. *Language Teaching Research*.

Boersma, P., & Weenink, D. (2013). *Praat: doing phonetics by computer*.

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.

Chambers, F. (1997). What do we mean by fluency? *System 25*(4): 535–544.

Council of Europe. (2011). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Council of Europe.

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

De Jong, N.H., & Bosker, H.R. (2013). Choosing a threshold for silent pauses to measure second language fluency. In R. Eklund (Ed.), *Proceedings of the 6th workshop on disfluency in spontaneous speech* (DiSS) (pp. 17–20). Royal Institute of Technology (KTH).

De Jong, N.H., Groenhout, R., Schoonen, R., & Hulstijn, J.H. (2015). Second language fluency: Speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics. 36*(2), 223–243.

De Jong N.H., Steinel, M.P., Florijn, A., Schoonen, R. & Hulstijn, J.H. (2012). The effect of task complexity on functional adequacy, fluency and lexical diversity in speaking performances of native and non-native speakers. In A. Housen, F. Kuiken, I. Vedder (Eds.) *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*. John Benjamins. 121–142.

Dewaele, J.-M. (2007). The effect of multilingualism, socio biographical, and situational factors on communication anxiety and foreign language anxiety of mature language learners. *The International Journal of Bilingualism, 11*, 391–409.

Eysenck, M.W., & Calvo, M.G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition and Emotion*, *6*(6), 409–434.

Eysenck, M.W., Santos Nazanin Derakshan, R. & Calvo. M.G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion 7*(2), 336–353.

Foster P., Tonkyn A., Wigglesworth G. (2000). Measuring spoken language: a unit for all reasons, *Applied Linguistics*, Volume 21, Issue 3, September 2000, Pages 354–375.

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

Hewitt, E., & Stephenson, J. (2012). Foreign language anxiety and oral exam performance: A replication of Phillips's MLJ study. *The Modern Language Journal*, *96*, 170–189.

Horwitz, E.K. (2013). *Becoming* a *language teacher*: *A practical guide* to *second language learning* and *teaching*, 2nd edition, Pearson.

Horwitz, E.K., Horwitz, B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*, ii, 125–132.

Horwitz, E.K. (2017). On the misreading of Horwitz, Horwitz and Cope (1986) and the need to balance anxiety research and the experiences of anxious language learners. In C. Gkonou, M. Daubney & J.-M. Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications* (pp. 31–47). Multilingual Matters.

Hunter, A-M. (2017). *Fluency development in the ESL classroom: The impact of immediate task repetition and procedural repetition on learners' oral fluency*. (Unpublished doctoral dissertation). University of Surrey.

Huensch, A., & Tracy–Ventura, N. (2016). Understanding second language fluency behaviour: The effects of individual differences in first language fluency, cross-linguistic differences, and proficiency over time. *Applied Psycholinguistics*, *38*, 755–785.

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

Kahng, J. (2014). Exploring utterance and cognitive fluency of L1 and L2 English speakers: Temporal measures and stimulated recall. *Language Learning*, 64, 809–854.

Kahng, J. (2022). Fluency. In: In T.M. Derwing, M.J. Munro & R.I. Thomson (Eds.), *The Routledge handbook of second language acquisition and speaking* (pp. 188–200). New York: Routledge.

Kormos, J. (1999). Monitoring and self-repair in L2. *Language Learning*, *49*, 303–342.

Kormos, J. (2000). The role of attention in monitoring second language speech production. *Language Learning 50*(2), 343–384.

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

Kormos, J. (2015). Individual differences in second language speech production. In J.W. Schwieter (Ed.), *The Cambridge handbook of bilingual processing* (pp. 369–388). Cambridge University Press.

Lambert, C., Kormos, J. & Minn, D. (2017). Task repetition and second language speech processing. *Studies in Second Language Acquisition*, vol. 39, no. 1, pp. 167–196.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behaviour Research Methods*, *44*(2), 325–343.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, *40*(3), 387–417.

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

Levelt, W.J.M., Roelofs, A., & Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1–37.

MacIntyre, P., & Gardner, R. (1989). Anxiety and second language learning: Toward a theoretical clarification. *Language Learning, 39*, 251–275.

MacIntyre, P.D. & R.C. Gardner. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning* 44(2), 283–305.

McDonough, K. & Trofimovich, P. (2008). *Using priming methods in second language research*. London: Taylor & Francis.

Mak, B. (2011). An exploration of speaking-in-class anxiety with Chinese ESL learners. *System, 39*(2), 202–214.

Michel, M.C., Kuiken, F., & Vedder, I. (2007). The influence of complexity in monologic versus dialogic tasks in Dutch L2. *International Review of Applied Linguistics in Language Teaching*, *45*(3), 241–259.

Pérez Castillejo, S. (2019). The role of foreign language anxiety on L2 utterance fluency during a final exam. *Language Testing*, 1–19.

Perez Castillejo, S. (2021). Prior processing, foreign language classroom anxiety, and L2 fluency. *International Review of Applied Linguistics in Language Teaching*.

Phillips, E.M. (1992). The effects of language anxiety on students' oral test performance and attitudes. *The Modern Language Journal 76*(1), 14–26.

Plonsky, L., & Oswald, F.L. (2014). How Big Is "Big"? Interpreting Effect Sizes in L2 Research. *Language Learning, 64*(4), 878–912.

R Core Team. (2020). R: *A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing.

Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes 14*(4), 423–441.

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*(1), 27–57.

Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics*, *45*, 193–213.

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*, 593–617.

Sanaei, O., Zafarghandi, A.M., & Sabet, M.K. (2015). The effect of classroom anxiety on EFL learners' oral narratives fluency: The case of intermediate level students. *Theory and Practice in Language Studies*, *5*(7), 1390–1400.

Segalowitz, N. (2010). *Cognitive bases of L2 fluency*. Routledge.

Skehan, P. (2015). Limited attention capacity and cognition: Two hypotheses regarding second language performance on tasks. In M. Bygate (Ed.), *Domains and directions in the development of TBLT: A decade of 191 plenaries from the international conference* (pp. 123–156). John Benjamins.

Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.). *Cognition and second language instruction* (pp. 183–205). Cambridge University Press.

Tavakoli, P, Nakatsuhara, F., Hunter A-M. (2020). Aspects of fluency across assessed levels of speaking proficiency, *Modern Language Journal*, *104*(1), pp. 169–191.

Tavakoli, P. and Wright, C. (2020). *Second language speech fluency: from research to practice*. Cambridge University Press.

Appendix A: Speaking tasks

You will perform two speaking tasks. You will give a monologue for both tasks.

Simple task:

> Read the first task.
> *What did you do last weekend?*
> *Talk about the activities you did.*
> You have a maximum of two minutes. If you finish earlier, you can stop speaking.
> Do you understand the task? The task will be recorded.

Complex task:

> Read the assignment.
> *In the Dutch lesson, you discuss the question:*
> ***Smartphones in the classroom: to ban or to allow?***
> *What is your opinion? Also, provide arguments to convince the other students.*
> You have a maximum of two minutes. If you finish earlier, you can stop speaking.
> Do you understand the task? The task will be recorded.

Appendix B

Table B.1 Results of mixed effects modeling analyses on factors affecting number of silent pauses

Fluency measure	Fixed effects: Factor	*Estimate*	*SE*	*p*	Random effects: Factor	Variance
Number of silent pauses	Intercept	52.63	14.04	<0.001	Participant	69.11
	Proficiency	−0.30	0.22	0.17		
	Anxiety	0.10	0.12	0.40		
	Task type	−1.05	1.49	0.48		

Table B.2 Results of mixed effects modeling analyses on factors affecting phonation time ratio

Fluency measure	Fixed effects: Factor	*Estimate*	*SE*	*p*	Random effects: Factor	Variance
Phonation time ratio	Intercept	50.50	11.32	<0.001	Participant	47.54
	Proficiency	0.28	0.18	0.12		
	Anxiety	−0.05	0.09	0.57		
	Task type	1.40	1.13	0.22		

Appendix C: Model comparisons

Speech rate
Anova (model 1, model 2)
model1: speechratetotal1 ~ 1 + proficiency + anxiety + Task + (1 | ID)
model2: speechratetotal1 ~ 1 + proficiency + anxiety + Task + as.factor(School) + (1 | ID)

Model	AIC	Deviance	Chisq	df	*p*
Model 1	1042.7	1030.7			
Model 2	1044.4	1028.4	2.33	2	0.31

Number of pauses
Anova (model 1, model 2)
model1: numofpauminute1 ~ 1 + proficiency + anxiety + Task + (1 | ID)
model2: numofpauminute1 ~ 1 + proficiency + anxiety+ Task + as.factor(School) + (1 | ID)

Model	AIC	Deviance	Chisq	df	*p*
Model 1	936.80	924.80			
Model 2	940.36	924.36	0.44	2	0.80

Number of mid-clause pauses
Anova (model 1, model 2)
model 1: Numbermidclausepauzes1 ~ 1 + proficiency + anxiety + Task + (1 | ID)
model 2: Numbermidclausepauzes1 ~ 1 + proficiency+ anxiety + Task + as.factor(School) + (1 | ID)

Model	**AIC**	**Deviance**	**Chisq**	**df**	***p***
Model 1	716.45	704.45			
Model 2	718.74	702.74	1.71	2	0.42

Phonation time ratio
Anova (model 1, model 2)
model1: PTR1 ~ 1 + proficiency + anxiety + Task + (1 | ID)
model2: PTR1 ~ 1 + proficiency + anxiety + Task + as.factor(School) + (1 | ID)

Model	**AIC**	**Deviance**	**Chisq**	**df**	***p***
Model 1	813.10	801.10			
Model 2	815.31	799.31	1.79	2	0.41

Silent pauses per minute
Anova (model 1, model2)
Model 1: sil.pausesminute1 ~ 1 + proficiency + anxiety + Task + (1 | ID)
Model 2: sil.pausesminute1 ~ 1 + proficiency + anxiety + Task + as.factor(School) +(1 | ID)

Model	**AIC**	**Deviance**	**Chisq**	**df**	***p***
Model 1	869.51	857.51			
Model 2	873.02	857.02	0.491	2	0.78