Accepted version of *Multilingual speakers' L1, L2, and L3 fluency across languages: A study of Finnish, Swedish, and English* (Peltonen & Lintunen 2022)

Published version available: https://doi.org/10.18261/nordand.17.1.4

Full reference for the published version:

Peltonen, P. & Lintunen, P. (2022). Multilingual speakers' L1, L2, and L3 fluency across languages: A study of Finnish, Swedish, and English. *Nordand,* 17, 48–63.

Multilingual speakers' L1, L2, and L3 fluency across languages: A study of Finnish, Swedish, and English

Pauliina Peltonen & Pekka Lintunen

Abstract

The present study provides a multilingual perspective on speech fluency by examining fluency across first language (L1), second language (L2), and third language (L3) productions in Finnish, Swedish and English among L1 Finnish (Group 1, G1) and Finnish–Swedish bilingual (Group 2, G2) university students in Finland. The two research questions focused on differences in speech fluency across the three languages between the groups and correlations across speech fluency measures in the different languages. 90 speech samples in Finnish (L1), Swedish (L1/L3), and English (L2) from 30 participants were analyzed in the present study. The speech samples consist of short picture narrations based on comic strip prompts. The fluency analyses focused on temporal fluency and stalling mechanisms. The data were analyzed quantitatively with Mann-Whitney U-tests and Spearman's rank-order correlation coefficients. The results demonstrated minor differences between the groups in their Finnish (L1) and English (L2) productions, but a higher level of fluency in Swedish for G2 (their L1) than G1 (their L3). For G1, the correlations were strongest between their L1 Finnish and L2 English, whereas for G2, mostly moderate to strong correlations were found between the different language pairs. The results suggest connections in fluency across the languages in a multilingual speaker's repertoire, but the strength of the correlations varies depending on the participants' L1(s) and proficiency level in the additional languages. Based on the findings, further research on fluency among multilingual speakers is needed. The results have implications for L2 teaching and assessment.

Keywords: fluency, multilingualism, speech production, individual differences

1. Introduction

Speech fluency is a key construct in second language (L2) learning, teaching, and assessment. As such, it has been widely studied as a component of L2 oral proficiency. While research comparing first language (L1) and L2 fluency from the same speakers has emerged as a promising avenue of research (e.g., De Jong et al., 2015), studies comparing speech fluency across multiple languages in a speaker's repertoire are rare. Examining learners' speech fluency characteristics in several languages can thus provide novel insights into the connections across L1, L2, and L3 speech fluency. In particular, our study fills a gap in research by examining both L1 and L3 productions in Swedish—a language that has rarely been included in fluency studies. Along with Swedish, L1 Finnish and L2 English samples are studied, enabling cross-linguistic comparisons between typologically related (Swedish and English) and unrelated languages (Finnish and Swedish/English).

Approaching speech fluency from a multilingual perspective, the present study aims to examine speech fluency across Finnish, Swedish, and English among L1 Finnish (Group 1, G1) and Finnish–Swedish bilingual university students (Group 2, G2) in Finland. A unique aspect of the present study is that the participants in both groups provided samples in the same languages: Finnish (L1), Swedish (L3 in G1, L1 in G2), and English (L2).[1] The study focuses on the differences in fluency between the two groups and the correlations in fluency across the participants' speech samples in different languages. The study is part of a project "Fluency across Multilingual Speakers" (*MultiFluency*; funded by the Swedish Cultural Foundation in Finland). From the data set collected for the project, 90 monologue speech samples in Finnish, Swedish, and English from 30 participants were chosen for the present study. The speech fluency analyses focused on two main aspects: temporal fluency, relating to the speed of talk and pausing, and stalling mechanisms, capturing resources that can be used to maintain fluency by providing planning time and helping to avoid long silences (see Dörnyei & Kormos, 1998; Peltonen, 2020). The data were analyzed statistically with non-parametric Mann-Whitney U-tests and Spearman's rank-order correlations.

2. Theoretical background

The present study focuses on speech fluency, which is generally studied as an indicator of L2 (oral) proficiency (based on Lennon's 1990 narrow sense of fluency) along with accuracy and complexity (on the complexity-accuracy-fluency framework, see e.g., Housen et al., 2012). Fluency relates to the smoothness and effortlessness of (L2) speech and has traditionally been examined from speech samples based on temporal measurements, such as speech rate and pausing (e.g., Chambers, 1997; De Jong, 2016; Lennon, 1990). This utterance fluency approach (Segalowitz, 2010), involving the analysis of fluency-related features from speech samples, is also applied in the present study. The traditional, temporal perspective on fluency is complemented with an analysis of stalling mechanisms (Dörnyei & Kormos, 1998), such as filled pauses and filler words, that reveal how the speakers maintain the flow of speech with different resources and reduce time spent in silence (see also Peltonen, 2020).

While there is a long research tradition on L2 speech fluency, cross-linguistic differences in fluency have received relatively little attention compared to other factors. Yet, some fluency studies involving cross-linguistic comparisons (e.g., Raupach, 1980) suggest that temporal patterns in the L1 may be reflected in the L2 (e.g., pause duration, see Riazantseva, 2001). Considering the context of the present study, Finland, the differences between the Swedish-speaking minority and Finnish-

speaking majority in their spoken L2 English are of particular interest (for an early study comparing Finnish-speaking Finns', Swedish-speaking Finns', and Swedes' fluency in L2 English, see Lehtonen, 1979). The Swedish-speaking Finns are considered having an advantage over the Finnish-speaking Finns in the acquisition of L2 English due to the typological closeness between their L1 (Swedish) and L2 English, in particular, but also due to their knowledge of a typologically different language, Finnish, and, for many, Finnish–Swedish bilingualism (Ringbom, 2007). Yet, according to Ringbom (2007), the advantages for the Swedish-speaking Finns are likely to be greater at the beginning of their L2 studies, diminishing with increasing proficiency and potentially disappearing by university level. Peltonen and Lintunen (2016), who compared Finnish-speaking and Swedish-speaking learners' L2 English speech fluency at two proficiency levels, found that Swedish-speaking upper secondary school students indeed outperformed their Finnish-speaking peers in L2 English fluency, while the differences were minor at university level, supporting Ringbom's (2007) hypothesis. However, the study did not include L1 speech samples from the participants and the inferences about L1 influence were thus indirect, while the present study allows for more direct comparisons between L1 and L2 productions across two participant groups.

Compared to the study of cross-linguistic influences on L2 fluency, the role of general proficiency level in the L2 has been somewhat more widely studied, occasionally involving comparisons against a native speaker control group. The studies have demonstrated that L1 speech is typically faster (higher speech and/or articulation rate) and contains less pausing than L2 speech (e.g., Dumont, 2018; Götz, 2013; Kahng, 2014). The pauses are also more clearly concentrated at clause boundaries, while learner speech typically contains more mid-clause pauses (e.g., Kahng, 2014; Raupach, 1980). Gradually, with increasing proficiency, learners' fluency approaches native speaker fluency, although temporal measures, especially speed fluency, may show clearer development than other aspects (Kormos & Dénes, 2004; Peltonen & Lintunen, 2016; Tavakoli et al., 2020). While comparisons between learner and native speech fluency provide important information about how close learners are to the target norm in their L2 fluency, more recently, some studies have adopted alternative designs that enable within-subject comparisons (De Jong et al., 2015; Derwing et al., 2009; Duran-Karaoz & Tavakoli, 2020; Huensch & Tracy-Ventura, 2017; Peltonen, 2018). These studies have been motivated by the fact that also native speakers vary in their fluency (e.g., Götz, 2013; Kahng, 2014; Lennon, 1990) and L2 fluency features may thus, also, at least partly reflect the differences that underlie variation in L1 speech.

The studies examining the connections between L1 and L2 speech fluency have generally examined either one L1–L2 pair (e.g, Duran-Karaoz & Tavakoli, 2020; Peltonen, 2018) or two L1 groups studying the same L2, the latter design enabling the study of cross-linguistic influence (e.g., De Jong et al., 2015; Derwing et al., 2009). The accumulating evidence suggests that L2 fluency measures are, at least to some extent, correlated with L1 fluency measures and that L2 fluency behavior can be partially predicted based on L1 fluency measures. In an influential study exploring the effects of L1 speaking style on L2 fluency, De Jong et al. (2015) studied speakers of a typologically close L1 (English, $n = 27$) and typologically distant L1 (Turkish, $n = 24$) and their L2 Dutch productions. The correlational analyses (with the L1 groups examined together) showed positive, statistically significant, and mostly strong ($r > 0.6$) correlations between all seven L1 and L2 fluency measures. Regression analyses were in line with these findings, demonstrating that L1 fluency measures predicted some variance in all L2 fluency measures (ranging from 21% to 57%, combined with the L1 group information).

Other studies have provided additional support for the connections between L1 and L2 fluency, but also that the strength of the correlations may change with L2 proficiency development. Huensch and Tracy-Ventura (2017) examined the effects of study abroad on fluency with two groups of L1 English learners of Spanish ($n = 24$) and French ($n = 25$) and found moderate to strong positive correlations between L1 and L2 fluency before and after study abroad for both groups for two key temporal measures (mean syllable duration and silent pauses per second). However, the strength of the correlations varied between the groups for other measures of fluency: L2 French learners demonstrated more statistically significant and generally stronger correlations after the study abroad. The findings echo previous results from Derwing et al.'s (2009) seminal study on L2 English fluency development involving 16 Slavic L1 and 16 Mandaring L1 speakers. In the study, initial correlations between three L1 and L2 fluency measures (pauses per second, speech rate, and pruned speech rate) were found for both groups, but later only for the Slavic group. Both studies highlight shifts in L1–L2 correlations over time and suggest the relationship between L1 and L2 fluency may change with developing proficiency.

Some additional support that L2 proficiency level influences the strength of the L1–L2 fluency correlations has been provided in recent cross-sectional fluency studies; yet the findings are somewhat mixed. A study comparing L1 Finnish learners of English at two school levels ($N = 42$) with 13 fluency measures found more statistically significant and stronger (mostly medium to large)

correlations between L1 and L2 fluency for second year upper secondary school students compared to ninth-grade students (Peltonen, 2018). The upper secondary school group represented, on average, the Common European Framework of Reference for Languages [CEFR] level B2, while the ninth-grade group represented level B1; the results thus suggest that learners seem to come closer to their L1 fluency profile as they become more advanced in their L2. In another cross-sectional fluency study of L1 Turkish speakers studying L2 English (N = 42; representing the CEFR levels A2, B1, and B2), Duran-Karaoz and Tavakoli (2020) found statistically significant correlations between L1 and L2 fluency measures for four of the total seven variables (breakdown and repair measures), ranging from r = .30 to r = .60 in effect. However, partial correlations that controlled for proficiency level did not significantly affect the strength of the correlations, suggesting that the L1–L2 relationship was not mediated by proficiency level. Based on the two studies, the role of proficiency level in influencing the relationship between L1 and L2 fluency remains somewhat inconclusive and warrants further investigation. In the present study, for the L1 Finnish group, the proficiency level in their L2 English and L3 Swedish is different, which allows us to compare the strength of the correlations between their L1 Finnish and the two additional languages.

3. Methodology

3.1 Research questions and hypotheses

Based on the literature review, our study extends previous fluency research by examining the roles of cross-linguistic influence, proficiency level, and individual differences among the language repertoire of multilingual speakers. The two research questions (RQs) were formulated as follows:

1) How does speech fluency in Finnish, Swedish, and English differ between G1 and G2?

2) To what extent are speech fluency measures correlated across the three languages for G1 and G2?

For RQ 1, the differences between groups regarding Swedish were of particular interest due to the status of the language varying between the groups (L3 in G1, L1 in G2). In addition, based on previous research, cross-linguistic differences were expected to result in some, albeit relatively minor, differences in L2 English across the two groups (cf. Peltonen & Lintunen, 2016; Ringbom, 2007). For RQ 2, based on previous studies examining L1 and L2 correlations (De Jong et al., 2015; Derwing et al., 2009; Duran-Karaoz & Tavakoli, 2020; Huensch & Tracy-Ventura, 2017; Peltonen, 2018), both typological differences and proficiency level were expected to influence the strength of

the associations along with potential individual differences. Thus, for G1, we hypothesized the correlations to be stronger between their L1 Finnish and L2 English than between L1 Finnish and L3 Swedish due to a lower proficiency level in Swedish. For G2, a specific hypothesis was not formulated due to the lack of previous similar fluency studies involving bilingual participants. Finally, as a similar design involving samples in three languages from the same participants has not, to our knowledge, been used before, we did not have specific predictions regarding the potential individual speaking style influence across multiple languages.

3.2 Participants

The participants were chosen among students who participated in the *MultiFluency* project. Participation in the project was voluntary, and signed informed consent was obtained from the participants. From the project participants, 20 L1 Finnish participants (G1) were included in the present study. The main criterion for the inclusion was the reliability of the L3 Swedish sample: the L1 Finnish participants who did not provide samples in Swedish or provided samples that were too short for reliable analyses (less than 20 seconds in duration or containing fewer than 30 syllables) were excluded from the sample. To keep the sample as homogenous as possible, one additional L1 Finnish participant was excluded due to slight speech impairment. All the Finnish–Swedish bilingual project participants (n = 10; G2) were included in the present study.

All 30 participants were students of language subjects at two universities in Southern Finland. Based on the information provided in the background questionnaire, the participants in G1 were on average 20.60 years old (17 female, 2 male, 1 other) and 25.50 years old (8 female, 2 male) in G2. All participants in G1 reported Finnish as their native language. In G2, the participants identified themselves as Finnish–Swedish bilinguals (n = 6), reported Finnish as their L1 and Swedish as an additional language spoken at home (n = 3), or reported Swedish as their L1 (n = 1), but all used both Swedish and Finnish in their everyday lives. Considering that the data were collected in Southwest Finland, this is representative of the population. Furthermore, one participant from G1 and two from G2 reported English as an additional language spoken at home.

Regarding the participants' additional languages (L3 Swedish and L2 English for G1, L2 English for G2), G1 had studied Swedish at school for an average of 6.63 years. Three G1 participants studied Swedish at university (two as a major subject, one as a minor subject). Based on the average

of the four CEFR criteria for qualitative aspects of spoken language use (range, accuracy, fluency, and coherence; Council of Europe, 2001, pp. 28–29), assessments of the speech samples conducted by three trained raters suggested that G1 participants represented, on average, level A2 in spoken Swedish. Furthermore, G1 self-evaluated their speaking skills in Swedish on a scale from 1–5 on average as 1.85 (between 1 = weak and 2 = moderate). Regarding L2 English, both groups had studied English at school for an average of nine years. In G1, 16 participants studied English as their major and four as their minor subject. In G2, five participants studied English as their major, one as their minor, and four reported not studying English at university. Based on the averages of four CEFR criteria for spoken language, both groups were estimated to represent level C1 on average regarding their spoken English. Among both groups, the most commonly studied other languages were German, French, and Spanish.

3.3 Data collection and procedure

The data for the present study were collected during 2020–2021. On-site (G1; two G2 participants) and hybrid (eight G2 participants) data collection procedures were used due to the covid-19 pandemic. The on-site collection was organized in conjunction with contact teaching, and the tasks were administered in a language laboratory. The hybrid collection involved on-site collection individually with recorders for two participants; the rest participated online and were remotely recorded in Zoom. The participants first filled in a short background questionnaire (pen-and-paper or Webropol), followed by the speaking tasks, which involved producing short (1–2 minutes) monologue speech samples in Finnish, Swedish, and English based on a cartoon strip. Three different cartoon strips were used. The order of cartoons and languages was counterbalanced to control for potential order effects, with the exception that the first task was always done in Finnish in the on-site procedure and in Swedish in the hybrid procedure with the same cartoon (strip 1). In the on-site procedure, the participants completed the next task either in English or Swedish with strip 2 or 3, depending on the participant's testing group, followed by the final task in the remaining language. The same logic of alternating the languages (starting with Finnish or English) and cartoons (starting with strip 2 or 3) was followed in the hybrid procedure. The participants were instructed to tell a story in their own words based on the pictures in the cartoon strip and were given two minutes of planning time. They could look at the cartoon strip while speaking. Picture narrations are commonly used to elicit monologue data in speech fluency studies (e.g., Lennon, 1990), and two of the prompts used in the present study have been employed in previous projects (see e.g., Peltonen & Lintunen, 2016). The third prompt was chosen to be as comparable with the

other prompts as possible; all three comics consisted of six frames, did not include speech bubbles, and had a clear storyline.

The speech samples were transcribed and double-checked by two research assistants under the supervision of the first author. Final checks were made by the first author. The research assistants identified and measured the durations of silent pauses (SPs) with the speech analysis software Praat (Boersma & Weenink, 2020) using a script (De Jong & Wempe, 2009; SP cut-off point was set at 0.25 seconds, following De Jong & Bosker, 2013), followed by manual checks of the SP boundaries. The research assistants also annotated four types of stalling mechanisms in Praat. The SP and stalling mechanism annotations were double-checked by the first author, followed by manual calculations of syllables and annotations of SP location. The SP location annotation was based on distinguishing mid-clause pauses from clause boundary and The Analysis of Speech Unit (AS-unit) boundary pauses (Foster et al., 2000). To calculate the fluency measures, frequencies and durations of the annotated features were extracted with a script (Lennes, 2002).

3.4 Fluency measures

Based on previous research, 15 fluency measures were chosen for the present study. The measures have been widely used and found to be reliable in characterizing both learners' and native speakers' fluency (e.g., De Jong, 2016; De Jong et al., 2015; Götz, 2013; Huensch & Tracy-Ventura, 2017). Furthermore, they ensure comparability with previous studies focusing on the connections between L1 and L2 fluency. Two main aspects of fluency were examined, temporal fluency (measures 1–9) and stalling mechanisms (measures 10–15; see also Peltonen, 2020):

1. speech rate (SR; syllables per minute of total time),
2. articulation rate (AR; syllables per minute of speaking time, excluding SPs),
3. phonation-time ratio (PTR; total speaking time divided by the total sample length),
4. the number of SPs,
5. the number of mid-clause SPs,
6. the number of clause/AS-unit boundary SPs,
7. the mean length of SPs,
8. the mean length of mid-clause SPs,
9. the mean length of clause/AS-unit boundary SPs,
10. the number of filled pauses (FPs),

11. the mean length of FPs,
12. the number of drawls (sound elongations of 0.30 seconds or longer),
13. the mean length of drawls,
14. the number of fillers, and
15. the number of repetitions (words or longer stretches of speech repeated without modification).

All frequency measures were standardized per minute of speaking time (i.e., excluding SPs; De Jong, 2016). Temporal fluency included measures of *speed fluency* (AR), *breakdown fluency* or simply *pausing* (SP measures), and *composite* measures (SR, PTR; Skehan, 2009, 2014). The stalling mechanisms included FPs, fillers, drawls, and repetitions. FPs refer to non-lexicalized pauses (*uh, um, eh* in English/Swedish; *öö* in Finnish) and fillers to lexicalized pauses (Dörnyei & Kormos, 1998), including discourse markers *like, you know, well* (*elikkä/elikkäs, niinku/niiku, tota/tuota, no* in Finnish; *liksom, typ, nå* in Swedish) and smallwords *sort of, kind of* (*tämmönen/tällanen/tommonen/semmonen* in Finnish; *det här/det där* in Swedish) (Götz, 2013; Peltonen, 2018). The stalling mechanisms were expected to demonstrate individual variation in how speakers maintain fluency (e.g, Dumont, 2018; Götz, 2013; Peltonen, 2020; Wolk et al., 2020). As a novel aspect, the average durations of FPs and drawls were examined along with their frequencies.

3.5 Statistical analysis

The statistical analyses were performed in SPSS (version 26). Based on a visual examination of boxplots and the results of Shapiro-Wilk tests of normality, the non-parametric Mann-Whitney U tests were run to answer RQ 1, since some variables for one or both groups violated the assumptions of normality and/or included outliers. Effect sizes (r) were calculated to estimate the magnitude of the differences (Larson-Hall, 2010, pp. 377–378). Due to multiple comparisons across the two groups (15 fluency measures for each language sample), the False Discovery Rate (FDR) method was applied to control for potential type I errors (Larson-Hall, 2010, pp. 251–252). To answer RQ 2, Spearman's rank-order correlation coefficients were calculated (separately for the two groups). Plonsky and Oswald's (2014, p. 889) guidelines were used to interpret the effect sizes (r = 0.25 as small, r = 0.40 as medium, and r = 0.60 as large effect).

4. Results

There was some variation in the samples produced by the two groups, the L1 Finnish group (G1) producing on average somewhat shorter samples (Finnish M = 44.14 s, Swedish M = 56.75 s, English M = 53.86 s) than the Finnish–Swedish bilingual group (G2; Finnish M = 81.50 s, Swedish

M = 77.15 s, English *M* = 87.26 s). This may be due to differences in the data collection procedures, which involved simultaneous recordings in a language laboratory for G1 and individual recordings for the majority of G2. In line with the differences in the sample length, G1 produced, on average, also fewer syllables in all languages (Finnish *M* = 169.45, Swedish *M* = 82.90, English *M* = 139.30) compared to G2 (Finnish *M* = 328.40, Swedish *M* = 258.50, English *M* = 242.80).

4.1 Differences in fluency between Finnish-speakers and Finnish–Swedish bilinguals

Regarding RQ 1, G1 and G2 were compared for their speech fluency in three languages: Finnish (L1), Swedish (L1/L3), and English (L2). Descriptive statistics (*M*, *SD*) for the 15 fluency measures on each language are compiled in Table 1. To examine whether differences in fluency between the two groups were statistically significant, three sets of Mann-Whitney U tests (all fluency measures for each language) were calculated. The results are presented in Table 2, where statistically significant differences ($p < .05$) are marked in bold, with accompanying *r* effect sizes. Following the FDR correction, the differences in Swedish fluency for mean length of end-clause SPs ($p = .020$) and fillers per minute of speaking time ($p = .032$) were not considered statistically significant.

Table 1
Speech Fluency Measures for Finnish, Swedish, and English Productions by Group

	Finnish		Swedish		English	
	G1	G2	G1	G2	G1	G2
Measure	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Temporal fluency						
Speech rate	229.76 (47.03)	221.57 (51.70)	88.25 (24.68)	197.49 (41.09)	159.09 (26.71)	165.47 (18.39)
Articulation rate	324.77 (44.74)	305.55 (51.79)	177.78 (34.45)	275.66 (38.08)	225.04 (22.29)	224.13 (10.64)
Phonation-time ratio	0.70 (0.08)	0.72 (0.10)	0.50 (0.09)	0.71 (0.11)	0.71 (0.08)	0.74 (0.09)
Silent pauses (SPs) / min. of ST	33.83 (10.36)	30.57 (11.38)	60.02 (17.17)	31.90 (10.86)	36.41 (12.95)	30.01 (7.50)
Mid-clause SPs / min. of ST	12.95 (6.37)	14.48 (6.38)	34.42 (13.54)	12.60 (8.12)	18.02 (10.22)	13.08 (5.93)

Clause boundary SPs / min. of ST	20.88 (9.67)	16.09 (7.11)	25.61 (10.46)	19.29 (4.54)	18.39 (5.76)	16.93 (3.31)
Mean length of SPs (s)	0.77 (0.16)	0.78 (0.17)	1.08 (0.25)	0.77 (0.20)	0.72 (0.14)	0.71 (0.18)
Mean length of mid-clause SPs (s)	0.58 (0.18)	0.63 (0.14)	0.95 (0.25)	0.64 (0.17)	0.57 (0.14)	0.57 (0.14)
Mean length of end-clause SPs (s)	0.86 (0.22)	0.89 (0.30)	1.18 (0.35)	0.86 (0.28)	0.83 (0.17)	0.82 (0.25)
Stalling mechanisms						
Filled pauses (FPs) / min. of ST	3.16 (3.65)	3.79 (2.83)	18.79 (6.98)	6.21 (4.29)	5.59 (5.13)	8.02 (5.46)
Mean length of FPs (s)	0.24 (0.25)	0.33 (0.21)	0.46 (0.11)	0.29 (0.17)	0.29 (0.23)	0.32 (0.13)
Drawls / min. of ST	2.58 (3.83)	2.73 (2.18)	17.85 (11.50)	5.36 (4.34)	6.28 (4.27)	6.71 (3.99)
Mean length of drawls (s)	0.22 (0.25)	0.33 (0.19)	0.38 (0.15)	0.31 (0.17)	0.34 (0.12)	0.34 (0.13)
Fillers / min. of ST	2.54 (3.57)	3.99 (5.91)	0.13 (0.40)	1.80 (3.27)	0.41 (0.75)	2.45 (5.00)
Repetitions / min. of ST	0.57 (1.06)	2.67 (2.63)	3.46 (5.46)	2.62 (4.08)	2.75 (4.19)	1.19 (1.28)

Note. G1 = Group 1 (*n* = 20), G2 = Group 2 (*n* = 10); ST = speaking time.

Table 2

Results of the Mann-Whitney U Tests Comparing L1 Finnish and Finnish–Swedish bilinguals' Fluency in Finnish, Swedish, and English with Effect Sizes (*r*)

	Finnish			Swedish			English		
Measure	*Z*	*p*	*r*	*Z*	*p*	*r*	*Z*	*p*	*r*
Temporal fluency									
Speech rate	–.220	.826	.04	**–4.267**	**<.001**	**.78**	–.704	.481	.13
Articulation rate	–.704	.481	.13	**–4.179**	**<.001**	**.76**	–.660	.509	.12
Phonation-time ratio	–.880	.379	.16	**–3.871**	**<.001**	**.71**	–.880	.379	.16
Silent pauses (SPs) / min. of ST	–1.100	.271	.20	**–3.740**	**<.001**	**.68**	–1.232	.218	.23

Mid-clause SPs / min. of ST	–.176	.860	.03	**–3.652**	**<.001**	**.67**	–.924	.356	.17
Clause boundary SPs / min. of ST	–1.716	.086	.31	–1.892	.059	.35	–.748	.455	.14
Mean length of SPs (s)	–.000	1.000	.00	**–2.772**	**.006**	**.51**	–.088	.930	.02
Mean length of mid-clause SPs (s)	–.528	.598	.10	**–3.212**	**.001**	**.59**	–.044	.65	.01
Mean length of end-clause SPs (s)	–.044	.965	.01	–2.332	.020	.43	–.572	.567	.10
Stalling mechanisms									
Filled pauses (FPs) / min. of ST	–.857	.391	.16	**–3.784**	**<.001**	**.69**	–1.391	.164	.25
Mean length of FPs (s)	–1.128	.259	.21	**–2.948**	**.003**	**.54**	–.199	.842	.04
Drawls / min. of ST	–.909	.363	.17	**–2.863**	**.004**	**.53**	–.264	.792	.05
Mean length of drawls (s)	–1.455	.146	.27	–1.409	.159	.26	–.044	.965	.01
Fillers / min. of ST	–.812	.417	.15	–2.141	.032	.39	–.706	.480	.13
Repetitions / min. of ST	**–2.608**	**.009**	**.48**	–.505	.614	.09	–.773	.440	.14

The differences in Finnish speech fluency across the two groups were relatively minor based on the mean values (Table 1), and no statistically significant differences were detected for the temporal measures (Table 2). Regarding stalling mechanisms, the averages were also relatively close for the two groups; yet G2 produced somewhat more fillers and repetitions than G1 in Finnish. Based on the Mann-Whitney U tests and FDR corrections, only the difference in repetitions reached statistical significance (medium effect size). In addition to the group averages, the high SDs for some of the stalling mechanisms, especially fillers, indicate relatively strong individual variation in the use of these resources (Table 1). Overall, the results indicate that speech fluency in Finnish based on temporal measures and stalling mechanisms is comparable across the two groups despite Finnish being the only L1 for G1 participants and one of their L1s along with Swedish for the bilingual G2 participants.

When comparing the speech fluency in Swedish between the groups, clear differences were found: ten of the total 15 measures (seven temporal measures; three stalling mechanisms) differed statistically significantly between the groups with medium to large effect sizes. The results suggest that speech fluency in Swedish was higher for the Finnish–Swedish bilingual group (G2) compared to G1, who spoke Swedish as their L3 (Table 2). In particular, the large effect sizes associated with five temporal measures indicate that G1 spoke Swedish more slowly (based on both SR and AR), spent more time in silence (lower PTR), paused twice as frequently overall, and exhibited more mid-clause pausing. The statistically significant differences in three stalling mechanisms further suggest that G1 participants produced more drawls and more and longer FPs on average in Swedish than G2 participants.

Finally, regarding L2 English fluency, the differences between the groups were minor based on the averages (Table 1). The Mann-Whitney U tests confirmed that the differences in fluency were not statistically significant, and all effect sizes can be considered small (Table 2). While the averages for SR and AR for both groups were somewhat lower in L2 English than in L1 Finnish and L1 Swedish, several pausing measures were, on average, close to the L1 Finnish and L1 Swedish means. Both groups thus exhibited a high level of L2 English fluency based on the temporal measures, especially pausing. Regarding stalling mechanisms, both groups used FPs and drawls relatively often, while fillers and repetitions were less frequently used. As indicated by the high SDs, the use of stalling mechanisms seemed to be prone to individual variation also in L2 English.

4.2 Correlations in fluency across Finnish, Swedish, and English

The correlations in fluency across the three languages were examined separately for the two groups with Spearman rank-order correlations. The results for the L1 Finnish group (G1) are compiled in Table 3 and the results for the Finnish–Swedish bilingual group (G2) in Table 4.

Table 3
Correlations in Fluency across Languages (Finnish, Swedish, and English) for G1

Measure	G1 (n = 20): FIN–SWE (L1–L3)	G1 (n = 20): FIN–ENG (L1–L2)	G1 (n = 20): ENG–SWE (L2–L3)
Temporal fluency			

Speech rate	.060	.606**	.302
Articulation rate	.349	.768**	.415
Phonation-time ratio	–.078	.611**	.373
Silent pauses (SPs) / min. of speaking time (ST)	.066	.504*	.514*
Mid-clause SPs / min. of ST	.332	.271	.598**
Clause boundary SPs / min. of ST	–.050	.611**	.041
Mean length of SPs (s)	.161	.562**	.335
Mean length of mid-clause SPs (s)	.069	.114	.215
Mean length of end-clause SPs (s)	.448*	.534*	.056
Stalling mechanisms			
Filled pauses (FPs) / min. of ST	.240	.219	.395
Mean length of FPs	–.203	.030	.285
Drawls / min. of ST	.475*	.440	.631**
Mean length of drawls	.112	.362	.594**
Fillers / min. of ST	.393	.390	.525*
Repetitions / min. of ST	.013	.441	.571**

Note. * $p < .05$, ** $p < .01$.

Table 4

Correlations in Fluency across Languages (Finnish, Swedish, and English) for G2

Measure	G2 (n = 10): SWE–FIN (L1–L1)	G2 (n = 10): SWE–ENG (L1–L2)	G2 (n = 10): FIN–ENG (L1–L2)
Temporal fluency			
Speech rate	.333	.212	.661*
Articulation rate	.200	.006	.539
Phonation-time ratio	.685*	.552	.806**
Silent pauses (SPs) / min. of speaking time (ST)	.721*	.564	.697*
Mid-clause SPs / min. of ST	.770**	.588	.745*
Clause boundary SPs / min. of ST	.576	.491	.164
Mean length of SPs (s)	.406	.624	.770**

Mean length of mid-clause SPs (s)	–.042	.661*	.370
Mean length of end-clause SPs (s)	.564	.612	.830**
Stalling mechanisms			
Filled pauses (FPs) / min. of ST	.500	.498	.596
Mean length of FPs	.768**	.395	.383
Drawls / min. of ST	.780**	.328	.328
Mean length of drawls	.841**	.584	.523
Fillers / min. of ST	.747*	.529	.485
Repetitions / min. of ST	.831**	–.107	.076

Note. * $p < .05$, ** $p < .01$.

As Table 3 demonstrates, for G1, the correlations between L1 Finnish and L3 Swedish speech fluency were mostly weak and small in effect. Only two correlations reached statistical significance, the mean length of end-clause SPs and the number of drawls, with medium effect sizes. The correlations between L1 Finnish and L2 English speech fluency were stronger and the majority of the correlations for temporal fluency measures reached statistical significance. The correlations can be considered strong for SR, AR, PTR, and clause boundary SPs per minute, while the correlations for SPs per minute, and the mean length of SPs and end-clause SPs were moderate. For stalling mechanisms, the correlations for drawls per minute and repetitions per minute were also moderate, albeit not statistically significant. Finally, the correlations between L2 English and L3 Swedish for G1 displayed a somewhat mixed pattern. For temporal fluency, two measures, SP frequency and mid-clause SP frequency, were statistically significantly and moderately correlated, while for stalling mechanisms, all measures except the FP measures were found to be statistically significantly correlated, the effect sizes ranging from medium to large.

As Table 4 shows, of the three sets of correlations, most statistically significant correlations for G2 were found between their L1s, Swedish and Finnish. Three of the statistically significant correlations were demonstrated for temporal fluency measures: PTR, SP frequency, and mid-clause SP frequency (all large effect sizes). Correlations for three additional SP measures, clause boundary SPs and the mean length of SPs and end-clause SPs, can be regarded as medium in effect, despite not reaching statistical significance. In addition, all stalling mechanisms except the frequency of FPs correlated strongly and statistically significantly across Swedish and Finnish. Regarding the

correlations between the speakers L1s (Finnish and Swedish) and their L2 English, mostly moderate to strong positive correlations were detected. Somewhat surprisingly, however, only one correlation reached statistical significance between L1 Swedish and L2 English, namely the mean length of mid-clause SPs (strong effect). However, the correlations for most of the other temporal measures, excluding SR and AR, can also be considered moderate (PTR and the three SP frequency measures) or strong (mean length of SPs and end-clause SPs), despite not reaching statistical significance. For stalling mechanisms, no statistically significant correlations were detected, but three of the correlations were medium in effect (the frequency of FPs, the mean length of drawls, and the frequency of fillers). Similarly, the correlations between L1 Finnish and L2 English for G2 mostly ranged from moderate to strong, six temporal measures reaching statistical significance (all large in effect): SR, PTR, SP frequency, mid-clause SP frequency, mean length of SPs, and the mean length of end-clause SPs. Regarding stalling mechanisms, no statistically significant correlations were found, and the effect sizes were mostly medium.

5. Discussion

RQ 1 addressed the differences in speech fluency between the L1 Finnish (G1) and Finnish–Swedish bilingual (G2) groups across three languages in their repertoire: Finnish (L1), Swedish (L1/L3), and English (L2). The comparisons based on Mann-Whitney U tests, FDR corrections, and *r* effect size calculations revealed, overall, minor and statistically non-significant differences between the groups in their Finnish (L1) fluency and English (L2) fluency, the only exception being repetitions in Finnish: G2 produced more of them than G1 on average. In contrast, in line with our hypothesis, clear differences in Swedish fluency were found between the groups for most temporal fluency measures and three stalling mechanisms, with medium to large effects. The difference can be explained by the fact that while Swedish was the L1 for G2, the participants in G1 had a relatively low proficiency level in their L3 Swedish and many participants had not used the language in their everyday lives since their upper secondary school studies.

The fact that the two groups did not, on average, differ in their L2 English speech corroborates previous findings and hypotheses regarding differences between Finnish-speaking and Swedish-speaking Finns' (spoken) L2 English at advanced levels of proficiency. Notably, from the perspective of speech fluency, the results support Ringbom's (2007) hypothesis predicting that the initial advantage for Swedish-speaking Finns in acquiring L2 English disappears by university level. The findings are also in line with Peltonen and Lintunen (2016), who found that Swedish-

speaking Finns spoke L2 English more fluently than their Finnish-speaking peers in upper secondary school, but not at university level, where the differences in L2 English fluency were minor. Compared to Peltonen and Lintunen (2016), however, where all university level participants studied English, in the present study, four G2 participants did not study English at university level, while all G1 participants did. Therefore, some differences in L2 English fluency could have emerged between the two groups, as we hypothesized. However, it is possible that the oral proficiency in English was high enough among participants in both groups, as suggested by the CEFR level estimates, to complete the rather simple task used in the present study with relative ease; a more demanding task could have revealed more differences between the groups.

Concerning RQ 2 that addressed the correlations between fluency measures across the three languages, the results demonstrated differing patterns for the two groups. As expected, for G1, more statistically significant and stronger correlations were found between their L1 Finnish and L2 English than between their L1 Finnish and L3 Swedish fluency. These differences are likely due to the higher proficiency level in English compared to Swedish; that is, G1's L2 English fluency is closer to that of their native language fluency. This finding lends further support for previous findings suggesting that the strength of L1–L2 correlations may change as L2 proficiency increases (Derwing et al., 2009; Huensch & Tracy-Ventura, 2017; Peltonen, 2018; but cf. Duran-Karaoz & Tavakoli, 2020). In addition to the correlations between L1 Finnish and the two additional languages, some statistically significant, moderate to strong correlations were found between L2 English and L3 Swedish for G1, notably for stalling mechanisms. Thus, the resources individual speakers used to maintain their flow of speech—especially drawls, fillers, and repetitions—were connected across their additional languages but less strongly linked to their L1. In other words, the participants may prefer similar ways of maintaining fluency across languages, which may not necessarily relate to the resources they use in their L1 (see also Götz, 2013; Peltonen, 2018). While this finding provides initial insights into fluency-maintaining mechanisms across additional languages, a largely unexplored area in fluency research, further qualitative analyses focusing on the use of these resources are needed. Ideally, more extensive quantitative analyses utilizing larger sample sizes would also be conducted to confirm these preliminary findings, since especially fillers in L3 Swedish and repetitions in L1 Finnish (G1) appeared infrequently in the data.

For the Finnish–Swedish bilingual speakers (G2), more moderate to strong correlations were found between the different languages in their repertoire compared to G1. Of the three sets of comparisons,

most statistically significant correlations were detected between their two L1s, Swedish and Finnish (with large effect sizes). In addition, mostly moderate to strong correlations were found regarding both Finnish–English fluency correlations and Swedish–English fluency correlations. More statistically significant correlations were detected between Finnish and English than between Swedish and English, but this can be attributed to some extent to the small sample size, since most of the non-statistically significant correlations were still medium in effect. While the clearest associations were found between the two L1s for G2, especially regarding stalling mechanisms, the fluency profiles between the L1(s) and L2 English were also connected.

Overall, the results of the correlational analyses corroborate previous findings of mainly moderate to strong correlations between L1 and L2 fluency and weaker associations when the proficiency level in the additional language is lower, as in the case of Swedish for G1 (De Jong et al., 2015; Derwing et al., 2009; Duran-Karaoz & Tavakoli, 2020; Huensch & Tracy-Ventura, 2017; Peltonen, 2018). Furthermore, the findings extend previous research by demonstrating relatively strong links in speech fluency across L1s (for G2) and providing preliminary support for profiles in stalling mechanism use across L2 and L3 productions (for G1).

6. Conclusion

The aim of the present study was to characterize multilingual speakers' speech fluency across three languages: Finnish, Swedish, and English. While previous studies have examined fluency in various L1s and L2s, this study was among the first to incorporate speech samples in three languages from the same speakers, with the status of the languages varying across the two groups. Based on the results, differences in fluency seem to be shaped by the proficiency level, but the role of cross-linguistic influences is minor at advanced levels of proficiency. Furthermore, multilingual speakers' fluency patterns seem to be connected across the languages in their repertoire, but the strength of the correlations is moderated by the participants' L1 profile and proficiency level in the additional languages. However, it should be noted that the proficiency level estimates were not based on proficiency tests, but rather on CEFR assessments and the participants' background information (self-assessed skill level and the number of years studying the language at school, in particular), which can be considered a limitation of the present study. Furthermore, the generalizability of the findings is limited by the relatively small sample size, especially regarding G2. In future, to complement the findings regarding speech fluency in the present study, other aspects of proficiency, such as accuracy and complexity, could also be examined from the productions to reveal potential

trade-off effects (e.g., Skehan, 2009, 2014). For instance, due to the relatively low proficiency level in Swedish, G1 could have spent their attentional resources mostly on accuracy rather than fluency.

Despite the limitations outlined above, the present study has provided novel insights into L1, L2, and L3 speech fluency from a multilingual perspective. The design of the present study allowed for the simultaneous study of multiple factors influencing fluency, including cross-linguistic influence, proficiency level, and individual differences, and has thus highlighted the benefits of incorporating speech samples in different languages from the same speakers within a single study. Based on the initial findings provided in this exploratory study, extending fluency research to involve bilingual speakers (and speech samples in both L1s) as well as comparisons across multiple additional languages in a speakers' repertoire could provide novel insights into the nature of speech fluency and the factors influencing fluency. The multilingual approach to fluency introduced in the present study also has implications beyond fluency research: in L2 teaching, raising awareness of stalling mechanisms might help learners to employ these resources across their additional languages more efficiently to fill in silences and to maintain fluency (see also Lintunen et al., 2020; Peltonen, 2020). Furthermore, the results indicate that acknowledging potential individual differences in fluency is important in both L2 teaching and L2 fluency assessment. In future, extending the design of the present study by combining fluency assessments (perceived fluency) and fluency analyses (utterance fluency) across different languages could provide the next step in the study of multilingual speakers' fluency.

References

Boersma, P., & Weenink, D. (2020). *Praat: Doing phonetics by computer* (Version 6.1.09) [Computer software]. http://www.praat.org/

Chambers, F. (1997). What do we mean by fluency? *System, 25*(4), 535–544. https://doi.org/10.1016/S0346-251X(97)00046-8

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

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., & 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). https://www.isca-speech.org/archive/diss_2013/dis6_017.html

De Jong, N. H., Groenhout, R., Schoonen, R., & Hulstijn, J. J. (2015). Second language fluency: Speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics, 36*(2), 223–243. https://doi.org/10.1017/S0142716413000210

De Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods, 41*, 385–390. https://doi.org/10.3758/BRM.41.2.385

Derwing, T. M., Munro, M. J., Thomson, R. I., & Rossiter, M. J. (2009). The relationship between L1 fluency and L2 fluency development. *Studies in Second Language Acquisition, 31*(4), 533–557. https://doi.org/10.1017/S0272263109990015

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.

Duran-Karaoz, Z., & Tavakoli, P. (2020). Predicting L2 fluency from L1 fluency behavior: The case of L1 Turkish and L2 English speakers. *Studies in Second Language Acquisition, 42*(4), 671–695. https://doi.org/10.1017/S0272263119000755

Dörnyei, Z., & Kormos, J. (1998). Problem-solving mechanisms in L2 communication: A psycholinguistic perspective. *Studies in Second Language Acquisition, 20*(3), 349–385. https://doi.org/10.1017/S0272263198003039

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics, 21*(3), 354–375. https://doi.org/10.1093/applin/21.3.354

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

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). Understanding second language fluency behavior: The effects of individual differences in first language fluency, cross-linguistic differences, and proficiency over time. *Applied Psycholinguistics, 38*(4), 755–785. https://doi.org/10.1017/S0142716416000424

Kahng, J. (2014). Exploring utterance and cognitive fluency of L1 and L2 English speakers: Temporal measures and stimulated recall. *Language Learning, 64*(4), 809–854. https://doi.org/10.1111/lang.12084

Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System, 32*(2), 145–164. https://doi.org/10.1016/j.system.2004.01.001

Larson–Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. Routledge/Taylor & Francis.

Lehtonen, J. (1979). Speech rate and pauses in the English of Finns, Swedish-speaking Finns, and Swedes. In R. Palmberg (Ed.), *Perception and production of English: Papers on interlanguage* (pp. 35–51). Åbo Akademi University.

Lennes, M. (2002). *Total duration of labeled segments* [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*(3), 387–417. https://doi.org/10.1111/j.1467-1770.1990.tb00669.x

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.

Peltonen, P. (2018). Exploring connections between first and second language fluency: A mixed methods approach. *The Modern Language Journal, 102*(4), 676–692. https://doi.org/10.1111/modl.12516

Peltonen, P. (2020). *Individual and interactional speech fluency in L2 English from a problem-solving perspective: A mixed-methods approach*. University of Turku. https://urn.fi/URN:ISBN:978-951-29-8137-3

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*(2), 209–238. https://doi.org/10.1515/eujal-2014-0018

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878–912. https://doi.org/10.1111/lang.12079

Raupach, M. (1980). Temporal variables in first and second language speech production. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler* (pp. 263–270). Mouton.

Riazantseva, A. (2001). Second language proficiency and pausing: A study of Russian speakers of English. *Studies in Second Language Acquisition, 23*(4), 497–526. https://doi.org/10.1017/S027226310100403X

Ringbom, H. (2007). *Cross-linguistic similarity in foreign language learning*. Multilingual Matters.

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

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*(4), 510–532. https://doi.org/10.1093/applin/amp047

Skehan, P. (2014). The context for researching a processing perspective on task performance. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 1–26). John

Benjamins.

Tavakoli, P., Nakatsuhara, F., & Hunter, A.-M. (2020). Aspects of fluency across assessed levels of speaking proficiency. *The Modern Language Journal, 104*(1), 169–191. https://doi.org/10.1111/modl.12620

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. https://doi.org/10.1007/s41701-019-00072-x

[1] L1 refers to the native language(s): Finnish for G1, Swedish and Finnish for the majority of G2 (for details, see Participants). L2 refers to English for both groups, and L3 refers to Swedish for G1.