university of groningen

umcg

University of Groningen

Learning an L2 and L3 at the same time: Help or hinder?

Huang, Ting

DOI:
10.33612/diss.135925259

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2020

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Huang, T. (2020). *Learning an L2 and L3 at the same time: Help or hinder?* [Thesis fully internal (DIV), University of Groningen]. https://doi.org/10.33612/diss.135925259

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 10-11-2026

CHAPTER 2

Learning an L2 and L3 at the same time: help or hinder?

Abstract | There is quite a bit of evidence showing that the experience of learning an L2 will help in learning an L3, but as far as we know, very little research has investigated the possible impact of L3 learning on the already existing and still developing L2 system within the learner. According to Complex Dynamic Systems Theory (CDST), language development depends on limited resources. In simultaneous L2 and L3 (L2+L3) learners, these resources have to be used for learning two languages, reducing the resources available for L2 learning. This might lead, firstly, to a slower L2 development in L2+L3 learners than in L2 only learners, and secondly, to more variability in the L2 during the learning process. In this chapter, we traced the L2 English writing proficiency scores (both in terms of gains and variability) of two groups of L1 Chinese learners within one academic year. One group learnt English only (L2) and the other group learnt English and Russian simultaneously. Results show that the L2+L3 learners did not develop their L2 to a lesser extent than the L2 learners did, but they showed more variability over time in one sub-area (fluency) of L2 writing proficiency. The implications are discussed.[1]

[1] This chapter has been slightly adapted for this dissertation and is published as a paper in: Huang, T., Steinkrauss, R., & Verspoor, M. (2020). Learning an L2 and L3 at the same time : help or hinder ? *International Journal of Multilingualism*, 1–17. https://doi.org/10.1080/14790718.2020.1779726

2.1 Introduction

An increasing number of Chinese adults are learning more than one foreign language at university. English, a compulsory foreign language course that every student starts to take from the third year of primary school until the end of the second year of university, is still the main foreign language. However, many universities have recently implemented new foreign language programs in which students major not only in English (L2) but in an additional foreign language (L3) at the same time as well, for example Japanese, German, or Russian. Ever since these new programs have been implemented, the question has risen whether students who learn an additional foreign language learn English as well as those who take only English (Zhu, 2008).

As the literature review will show, there is quite a bit of evidence that an L2 will help in learning an L3, but as far as we know, no research has been done on the effect of L3 learning on L2 learning. Therefore, it is important to investigate how learning a new third language might impact the already existing and still developing L2 system within the learner (de Bot, 2012).

Second language development is a complex and dynamic process, which involves interconnected subsystems (de Bot & Larsen-Freeman, 2011). Learning more than one foreign language simultaneously is certainly even more complex. According to complex dynamic systems theory and cognitive approaches to SLA (cf Skehan & Foster, 1997) the development of any (sub)system is dependent on internal and external resources. At the same time, resources are limited; therefore, there may be competition among (sub)systems for resources. Learning an L2 and L3 simultaneously can thus be assumed to compete for both internal resources such as attention, motivation and memory and external resources such as time available to learn the two languages. For instance, previous studies have shown that learners may have different L2 and L3 motivation, and in many cases, the motivations for different languages interact and impact each other (Csizér & Dörnyei, 2005; Dörnyei et al., 2006; Henry, 2010; Wang & Zheng, 2019; Zheng et al., 2019). Learning a new L3, such as Russian, alongside English as L2 may thus slow down the further development of the L2 due to the competition for motivation and the limited time available. In other words, learning the L2 and L3 simultaneously might hinder the learning of the L2. To test this assumption, the present paper compares the development of L2 English writing during one academic year in two types of students: those who learn only English (L2) and those who learn English and Russian simultaneously (L2+L3).

2.2 Background literature

How learning an L2 impacts L3 learning has been investigated quite often in recent years. For example, Schepens, Van Der Slik and Van Hout (2016) conclude that bilingual learners had an advantage in L3 learning, no matter what their L1 was. Exploring L3 acquisition may enhance our understanding of the language acquisition processes and provides an opportunity to investigate how previous linguistic language knowledge can help with learning a new language (Hammarberg and Williams, 2009). In the same vein, Cenoz (2010) argues that L3 learning has some unique characteristics compared to L2 learning, one of those being a possible impact of the L2 on L3 learning. Quite a bit of research investigated these possible effects on L3 learning (e.g. Kassaian & Esmae'li, 2011; Grey, Sanz, Morgan-Short, & Ullman, 2018; Sanz, 2000). As Grey et al. (2018) point out, the better performance in some linguistic aspects of the bilingual learner in additional language learning may be the result of the advantages that bilinguals have compared to the monolinguals. Some examples are stronger metalinguistic awareness (e.g. Hofer & Jessner, 2016), more effective learning strategies (e.g. Afsharrad & Sadeghi Benis, 2017; Grenfell & Harris, 2015), and higher language aptitude (e.g. Eisenstein, 1980).

To summarize, the more language learning experience learners have, the more learner-based advantages they enjoy for learning an additional language, and the easier it is for them to achieve a higher language proficiency in that language. The focus of the studies so far has been on the effect of L2 on L3. However, it seems reasonable to assume that the L2 may in turn be affected by learning the L3. To investigate this direction of impact would heed de Bot's (2012) call to investigate "one of the basic questions" in multilingual studies (p. 91) how the learning of an additional language impacts the already existing language system.

When the L2 and L3 are both developing, learning the L3 and L2 simultaneously may result in interactions between the L2 and L3 systems or subsystems. According to Complex Dynamic Systems Theory (CDST) (de Bot, 2017; Larsen-Freeman, 2017), complex dynamic systems consist of interconnected subsystems, which means that the development in any single subsystem leads to the development in other interconnected subsystems, which ultimately results in the development of the whole complex dynamic system (de Bot & Larsen-Freeman, 2011).

The complete connectedness of the subsystems can be well understood through the CDST studies on the development in L2 complexity, accuracy, and fluency (CAF). Different subsystems of the L2 system develop at a different rate, and the development of one subsystem might

happen only with the prerequisite of the development in another subsystem, such as lexical complexity and syntactical complexity (Verspoor et al., 2012). Different subsystems sometimes grow together such as word complexity and noun phrase complexity (Spoelman & Verspoor, 2010), and the complexity, accuracy and fluency in L2 speaking (Vercellotti (2017).

Such dynamic interactions may also apply to simultaneous L2 and L3 learning. Such a connectedness of linguistic subsystems has been amply attested in studies on cross-linguistic interference phenomena such as transfer (e.g. Murphy, 2005; Bardel & Falk, 2007) and attrition (e.g. Schmid & Köpke, 2017), and psycholinguistic phenomena such as lexical access (e.g. Van Heuven, Dijkstra, & Grainger, 1998; Jared & Kroll, 2001). In essence, "Languages that are frequently used cannot be completely deactivated" (Hermans, Bongaerts, De Bot, & Schreuder, 1998. p. 225).

At the same time, the dynamic process of development is dependent on resources, and resources accessible to humans in skill learning or problem solving are limited (Van Geert, 2003). Limited resources can generally be divided into internal resources and external resources. Internal resources are resources within the learning individuals, and in the case of language learning they may include the learner-based factors such as metalinguistic awareness or working memory, but also the capacity and time to learn, motivation, as well as already existing knowledge. External resources are resources outside the learning individuals, such as the social context, the time invested by the external environment, or the amount, quality, and type of linguistic input and instruction provided (de Bot & Larsen-Freeman, 2011).

A consequence of the resources being limited is that different subsystems may at times compete for resources, leading to a competitive relationship. Competitive subsystems develop in alternating patterns, which means that while one goes up, the other goes down (de Bot & Larsen-Freeman, 2011). Such competitive subsystems have been found in, for example, the lexicon versus syntax (e.g. Robinson & Mervis, 1998; Verspoor, Lowie, & Dijk, 2008), and noun phrase complexity versus sentence complexity (e.g. Spoelman & Verspoor, 2010). As the learner's linguistic proficiency increases and subsystems become automated and coordinated, they are less demanding on resources, and the competitive relation may change into a supportive or neutral one, as reported in Verspoor, Lowie, and Van Dijk (2008). Another effect of the limitation of resources, and the resulting competition, may be that an already existing system can be destabilized when another, new system is developing. Thelen and Smith (1994) specifically point to the fact that learning new behaviour may go temporarily at the expense of formerly established

or automated behaviour. In a similar vein, Kahneman (1973) argues that our automated behaviour (System 1) suffers somewhat when we have to consciously pay attention to something else (System 2).This (temporary) destabilization of a system is manifested in a higher degree of variability.

As an inherent property of a dynamic system, variability provides insight into the developmental process (de Bot, Lowie, & Verspoor, 2007; Lowie & Verspoor, 2019; Spoelman & Verspoor, 2010; van Geert & van Dijk, 2002). From a behavioral science view, variability is a result of both the external factors and the self-organization within the system itself, and can be seen as a predictor of development (de Weerth et al., 1999). For example, in Thelen and Smith's (1994) study, when the infant was ready to move to a new phase in motor development (phase shift), there was a higher degree of variability in the performance in the previous phase, meaning that before the stabilization of a certain movement, "the child may then explore a richer and more varied landscape of possible adaptive solutions" (p.88).

Assuming that the development of the L2 and the L3 systems are interconnected and relies on the availability of limited resources, we can expect some effect on both the L2 and L3. Arguably, learning two foreign languages simultaneously will require more resources than learning one. For instance, learning two languages simultaneously costs more time for the individuals, requires a relatively higher level of motivation, and needs more support from the external environment, like a more adequate language environment for both languages. As the amount of time and cognitive capacities available to a person are finite at any given moment, and since even a highly motivated L2+L3 learner may allocate more resources to the language they think should be their major priority, the other language would suffer in terms of resources available for learning it. With the L3 being completely new to the learners, it is likely to consume more time and attention than the L2.

Therefore, we may assume that the investment in an L3 may detract — at least temporarily — from the performance in the L2, as the two linguistic subsystems are competing for resources, resulting in lower L2 proficiency gains and a greater amount of variability than if only the L2 was learnt. The present study thus sets out to investigate the effect of simultaneous L2 and L3 learning on L2 learning. The following research questions guide this investigation:

(1) Does learning an L2 and L3 simultaneously lead to lower gains in L2 proficiency than when learning only the L2?

(2) Does learning an L2 and L3 simultaneously lead to more variability/destabilization in L2 proficiency than when learning only the L2?

Based on the argument of competing resources and the connectedness between the developing subsystems, we assume that simultaneously L2 and L3 learning will at least temporarily hinder L2 learning. Such a hindering effect may be shown in the language learning gains after a certain period (RQ 1) as well as in the variability of the learning process during this period (RQ 2).

2.3 Methodology

This longitudinal study examines the writing development of 72 L1 Chinese adult university foreign language majors in two programs and two cohorts, the first and second year. The two programs were a traditional foreign language program (English as major) and a new type of bi-foreign-language program (English and Russian as double major) at university level. English majors take English only in the first year, but in their second year they take another foreign language as minor, so in effect they then become L2+L3 learners too.

As this study is interested both in the L2 proficiency gains as well as the variability in L2 proficiency during the academic year, both a pre-post design and a longitudinal design was adopted. Pre-test scores (the mean of the first two writing samples) and post-test scores (the mean of the last two writing samples) were compared to investigate whether there was any difference between the two groups in gain scores. The variability in proficiency scores during the whole academic year was studied to assess whether the learning process of the two groups was different.

2.3.1 Participants

Of the 72 participants in the current study, 34 were English majors (L2 learners) from the traditional foreign language program, and 38 were English and Russian majors (L2+L3 learners) from the bi-foreign language program. For all learners, English was their L2, and for the English-Russian majors, Russian was their L3. All of the participants had been learning L2 English for about 10 years in school, and were at about an estimated B1 CEFR (Common European Framework of Reference) in reading and writing skills, while the English-Russian majors started to learn Russian only by the time they entered the university. The current study involved two cohorts, first year and second year students. Since the second-year L2 students in the traditional

program also started to learn a new L3 to a very limited extent, they were labelled L2+ students. See group sizes and instruction time in Table 2. 1.

Table 2.1: Groups and amount of language instruction time

	First year		Second year	
Groups	L2	L2+L3	L2+	L2+L3
Group size	23	22	11	16
Instruction (h/week)	E: 16	E:16 R:8	E:12 FL:4	E:12 R:10

Note. E: L2 English; R: L3 Russian, FL: L3 Japanese, Russian, German or French.

In order to be enrolled to the language major at this specific university, the applicants need to have achieved a relatively high score in the China National Higher Education Entrance Examination, and also meet a cut-off score of the English test. Therefore, upon entering the university, all of the participants started at a similar academic aptitude, and rather comparable English proficiency level. To further make sure that the L2+L3 and L2 learners started with a comparable English writing proficiency, their starting L2 writing proficiency were compared. The starting L2 writing proficiency was operationalized as the average total scores of the first two writing samples. The results showed that there was no significant difference between the L2+L3 learners and L2 learners, neither among the first year students ($t(43)=1.22$, $p=0.23$ (two-tailed)), nor the second year students ($t(25)=0.43$, $p=0.67$ (two-tailed)). Moreover, these groups had the same English teacher within each cohort. Thus proficiency level, academic aptitude and English teachers were controlled for.

2.3.2 Data collection procedure

Writing data were collected from the participants to assess their L2 proficiency. As a subtest of language proficiency, free writing data are suitable to assess both language-specific abilities such as the vocabulary size and the use of syntactic structures, and writing-specific abilities (Wolfe-Quintero, Inagaki, & Kim, 1998) such as composition. Because written texts show a language learners' active language use in all its facets, including the use of vocabulary, the knowledge about L2 idioms, tenses, errors and sentence constructions, Verspoor, Schmid and Xu (2012) argue that assessing writing samples is a particularly valuable way to measure general L2 proficiency.

The writing data were collected once every three weeks during one academic year (9 months). The writing tasks of the sample texts were the writing assignments for each unit of their

comprehensive English course. The total dataset of the present study consists of 810 writing samples. See Table 2. 2.

Table 2.2: The dataset

Numbers of participants and texts	First year		Second year	
Language groups	L2	L2+L3	L2+	L2+L3
Number of data points (during 9 months)	12		10	
Number of participants	23	22	11	16
Number of writing texts	276	264	110	160

2.3.3 Holistic rating rubric

As holistic proficiency rating scores are more generalizable than analytic scores (Schoonen, 2005), the current study adopted a holistic rating method. Second language proficiency is multi-componential in nature rather than unitary (Housen & Kuiken, 2009); therefore, to measure L2 performance adequately, comprehensive measures are needed. We chose to use holistic ratings for complexity, accuracy and fluency as a basis. Complexity, accuracy and fluency are a frequently used triad in L2 studies to capture learners' L2 performance and proficiency (Ellis & Barkhuizen, 2005; Housen, Kuiken, & Vedder, 2012; Skehan, 2009; Wolfe-Quintero et al., 1998). This includes CDST studies such as Spoelman and Verspoor (2010), Verspoor et al. (2012), Polat and Kim (2014), Chan, Verspoor, and Vahtrick (2015), Verspoor, Lowie, Chan, and Vahtrick (2017), and Lowie and Verspoor (2019).

Complexity is commonly described as the ability or willingness to produce more elaborate language using a varied range of sophisticated structures and vocabulary in the L2. The complexity measures in the present paper mainly focus on sentence structure. Accuracy is typically characterized as the degree of target-like and error-free use of language compared to the norm of the target language. The accuracy measures in the present study mainly focus on the degree of correctness of grammar use. Fluency in writing performance is often measured by the length of the text that the writer can produce within a certain duration of time as it reveals the degree of ease in writing (Wolfe-Quintero et al., 1998). In the present study, fluency was operationalized as text length (98 words and 758 words for the shortest and longest text respectively). To make the fluency score comparable to the other scores, fluency was recalculated as a score from 1 to 5 using the required minimum text length as a basis. As the course

progressed, this required text length increased from unit to unit. A text meeting the required length received a score of 2.75, and for every 30 words above or below this required text length, an additional score of 0.25 was added or subtracted.

As adequate and comprehensive as the CAF triad is, it fails to capture idiomaticity and coherence, which are also important indices to general writing proficiency, especially at higher levels of L2 writing (Hou et al., 2016). Idiomaticity in the present study is not to be equated with 'idioms'. While idioms are conventionalized combinations of words with figurative and often metaphorical meanings (Hou et al., 2016; Prodromou, 2003), in our use of the term idiomaticity, we include "conventionalized ways of saying things (CWOSTs)" (Smiskova, Verspoor, & Lowie, 2012, p.125), which are combinations of words in authentic and native-like ways. According to Prodromou (2003), idiomaticity is a necessary criterion to gauge leaner language because idiomaticity "makes language more real" (p.44), but it also is difficult even for the advanced students to acquire and therefore a useful criterion to assess language learners on. Finally, the texts were rated holistically on coherence as well, as Crossley and McNamara (2011) argued that coherence is another important aspect of writing proficiency. Coherence refers to the logical bonding and thematic development of the text and is also difficult to acquire (Hou et al., 2016; Kuo Chih-Hua, 1995).

Following the complexity, accuracy, fluency, idiomaticity and coherence (CAFIC) model by Hou et al. (2016), the present study developed its holistic rating rubrics in two steps. The first step was to build the rating rubrics based on an on-site grading and discussion of nine randomly selected texts from the sample. The rubric was thus specifically designed for the current dataset, making subtle differences possible within a general intermediate proficiency level. Details can be found in the Appendix A. The second step was to validate the holistic rubric developed in the first step. The nine texts were rated again on each aspect of CAFIC, using the 1-5 scale built in the previous step. A correlation analysis was carried out to investigate the relation between the initially chosen proficiency order in the first step and the total proficiency score based on the CAFIC measures of the 9 texts. Correlation analysis showed that the two ratings were significantly correlated ($r=0{,}826^{**}$, $p<0{,}01$), with a large effect size ($r^2=0.68$). This result suggests that the initial scores given by the three raters and the total scores based on the five CAFIC rubrics are highly in line with each other, validating the CAFIC holistic rubrics.

2.3.4 Rating procedure and intra-rater reliability

The remainder of the texts in the dataset were rated by the first author, who rated all texts of one unit (i.e. all texts collected at the same time, on the same topic) in one session before moving on to rate the texts of the next unit. To ensure the within-rater reliability, two methods were used during and after the rating process. To guarantee a consistent judgment from one unit to the next, ten texts randomly selected from the former unit were rerated every time before starting to rate the next unit. Then, the scores from the first and the second rating were compared. The rater continued to rate the texts from the next unit only once the two rating scores were highly and significantly correlated.

A further analysis was carried out to ensure the rater reliability during the whole rating process after finishing all the ratings. 50 texts were randomly selected by the second author from the whole dataset and rated by the first author a second time. The total scores (i.e. the sum of all five CAFIC) from the first rating and second rating were tested for correlations, and they were significantly correlated with large effect size, suggesting a within-rater reliability and consistency throughout the rating process. See Table 2. 3 for the details of the analysis on the 50 rerated texts.

Table 2.3: Correlation between the first and second rating of the 50 texts

Test	Complexity	Accuracy	Idiomaticity	Coherence	Holistic
Pearson's R	r=.713***	r=.772***	r=.587***	r=.610***	r=.749***

Note: ***, correlations are significant at 0.001 level

2.4 Analysis

All analyses were carried out for each CAFIC sub-score as well as for the total proficiency scores. The score for each CAFIC sub-measure varies from 1 to 5, and the total proficiency scores vary from 5 to 25. Pre-scores were the mean scores of the first two data points, and post-scores were the mean scores of the last two data points. Gain scores were the difference between the pre- and post-scores. Variability was operationalized as the coefficient of variation (CoV), calculated as the standard deviation divided by the mean of a certain period. Because we expected more destabilization early on, we tested the first half, the second half, and the whole period of the academic year for variability. The first period equals the first six (for the first-year students) or five (for the second-year students) data points, and the second half period equals the last six (for the first-year students) or five (for the second-year students) data points.

2.5 Results

For RQ 1, an independent-samples t-test was used to compare the gain scores of the L2 and L2+L3 students. For the second research question an independent-samples t-test was used, comparing the CoV of the L2 and L2+L3 learners. Any difference between L2+L3 and L2 learners would suggest that simultaneous L2 and L3 learning would have an impact on the L2 learning.

2.5.1 Gain scores in each group

The descriptive statistics show that for the first year students, the L2 learners increased from an average total score of 13.53 to 15.94, while the L2+L3 learners increased from 13.61 to 16.32. For the second year students, the L2 learners increased from 14.06 to 16.14, while the L2+L3 learners increased from 14.20 to 16.03. Based on these scores, the gain score for each group was calculated. An independent-samples t-test did not reveal any significant differences in the total L2 proficiency gains between the groups. But when investigating the CAFIC sub-scores individually, a significant difference was found in L2 complexity gain scores among second year students, with the L2+ learners (M=0.74, SE=0.07) gaining more in complexity (from an average score of 2.99 to 3.70) than the L2+L3 learners (M=0.38, SE=0.10), whose average score of complexity increased from 2.95 to 3.46, with this difference between gain scores being significant ($t(24.22) = 2.88$, $p = 0.007$). Tables 2. 4 and 2. 5 show the results.

Table 2.4: The differences on total proficiency gain scores among two groups

Total gain score				Independent sample t test	
		N	M (SE)	t(43)	r^2
First year	L2	23	2.61(0.32)	0.19	/
	L2+L3	22	2.70(0.34)		
Second year	L2+	11	2.03(0.32)	1.02	/
	L2+L3	16	1.60(0.27)		

Table 2.5: The differences in gain scores in complexity between two groups among second year students

Gain score in complexity				Independent sample t test	
		N	M (SE)	t(24.22)	r^2
Second year	L2+	11	0.74(0.07)	2.88**	0.26
	L2+L3	16	0.38(0.10)		

Note: **, difference is significant at 0.01 level.

2.5.2 Variability in the developmental process in two groups

The variability as measured by the CoV of the total proficiency scores as well as of each individual CAFIC sub-scores was analysed for the whole year as well as for the first half and the second half of the period of observation separately. The results showed that, 1) significant differences were only found among the first year students, not among second year students; 2) L2+L3 learners had a higher CoV than the L2 learners in terms of fluency in all three periods analysed, with a larger effect size for the first half (r^2=0.33) than the second half (r^2=0.13). Table 2. 6 shows the details.

Table 2. 6: Difference between first year L2+L3 vs. L2 learners in terms of fluency variability

Variability in Fluency				Independent sample t test	
		N	M (SE)	t(43)	r^2
First half	L2	23	0.06(0.006)	4.65***	0.33
	L2+L3	22	0.13(0.013)		
Second half	L2	23	0.10(0.006)	2.57***	0.13
	L2+L3	22	0.13(0.009)		
Whole year	L2	23	0.10(0.006)	3.76***	0.25
	L2+L3	22	0.14(0.010)		

Note: ***, difference is significant at 0.001 level.

2.6 Discussion

This study sought to investigate whether learning an L2 and L3 simultaneously helps or hinders the learning of the L2 by comparing two groups of learners, L2 learners and L2+L3 learners, on their L2 English writing proficiency as measured with holistic CAFIC rating scores. The results revealed that the two learner groups differed neither in total proficiency gain scores nor in any CAFIC sub-measure gain scores. An exception were the second-year students, where the L2+ learners outperformed the L2 +3 learners in L2 complexity gains. In addition, the two groups of first-year students differed significantly in variability regarding fluency. The implications are discussed in the following.

2.6.1 Improvement in CAFIC holistic scores in the two groups

We had assumed that resources for learning are limited and that learning an L2 and L3 simultaneously might place a higher demand on the learners' internal (e.g. cognition and motivation) and external (e.g. time) resources compared to learning one language only. Thus, we hypothesized that the group of simultaneous L2+L3 learners would gain less in L2 writing proficiency than the group of L2 only learners. Unexpectedly, this hypothesis was rejected. This result may suggest that the challenge of learning an L2 and L3 simultaneously does not necessarily hinder the development of the L2 regarding L2 writing proficiency in free writing tasks.

One possible reason for this could be that, although simultaneous L2+L3 learners may encounter time and motivation limitations for L2 learning, compared to L2 learners, they may enjoy other language learner-based advantages (internal resources), such as higher metalinguistic awareness, larger working memory capacity, and well-trained learning strategies, which could

compensate for the shortage of time available and limited motivation. Such a learner-based advantage was partly testified for the same population in another two studies within the same project concerning language learning motivation and working memory. With regards to motivation, the L2+L3 learners were found to have a significant higher motivation to learn L2 (English) than the L2 only learners (Huang, Steinkrauss, Verspoor, 2020a). The other study showed that while the L2+L3 learners and L2 learners started with a similar working memory capacity, the former had a significantly larger capacity than the latter after one academic year (Huang, Loerts & Steinkrauss, 2020). Therefore, the learner-based advantages among bilingual learners for L3 learning, such as learning strategies (e.g. Afsharrad & Sadeghi Benis, 2017; Grenfell & Harris, 2015), meta-linguistic awareness (e.g. Hofer & Jessner, 2016), and language aptitude (e.g. Eisenstein, 1980), may also be applicable to the L2 learning of L2+L3 simultaneous learners.

When exploring the data further, we found that second year L2+ learners gained more than the L2+L3 learners in one CAFIC sub-measure, i.e. complexity. Complexity in the current study refers to how simple and complex the sentence structures are in terms of dependent clauses or elaborated constructions. An example of low complexity from our writing samples is "They never complain about their life condition. Instead, they are being positive to their life". The characteristic of a low complexity text is that it mainly uses simple sentences, while a text of high complexity is featured as using many elaborated constructions; for example, one student wrote, "As is indicated from the wife's words, no matter how different the two's living background or any others are, the interracial marriage is available as long as there exists true love". See Appendix B for excerpts with different levels of complexity.

L2+ learners were a special group of learners in the present study: the L2+ learners had been L2 learners in their first year, but started to learn an additional L3 (Japanese, Russian, German or French) in the second year. However, this group had only 4 hours (versus 8-10 hours in the L2+L3 group) of L3 instruction per week. This finding is potentially in line with the idea that intensive simultaneous L2+L3 learning (as found in the L2+L3 learners) hinders L2 learning: the less intensive simultaneous learning experience of the L2+ leaners led to more L2 gains of that group. However, if this explanation were to apply, we would expect an even greater effect of this kind in our first-year learners, where the L2 learners were able to focus solely on learning the L2 and should consequently also have shown larger L2 complexity gains than L2+L3 learners. This, however, is not supported by the data as no difference in L2 gains on any CAFIC sub-measure

was found among the first-year students. Any attempt to explain this differential effect within the present study must therefore remain speculative. But what we can infer from the results is that compared to L2+L3 learners who had been learning the L3 for two years by the end of the study, the fact that the L2+ learners gained more in complexity would suggest that learning an L3 along with L2 does not hinder the learning of L2 in the early stage of learning the L3.

2.6.2 Different degrees of variability between L2+L3 and L2 learners regarding fluency

While the pre-test and post-test design revealed that learning two foreign languages at the same time did not lead them to gain less than the L2 learners, it did not provide any information on the process L2 development. To gain more insight into the developmental process, a variability analysis was carried out to see whether L2+L3 learners' L2 proficiency scores varied more than those of the L2 learners, as the patterns of variability reveal information about development (Lowie & Verspoor, 2015). The greater variability in L2 fluency among the first year L2+L3 learners suggested that the simultaneous L2 and L3 learning makes the developing L2 system more unstable than in the L2 only learners. The results further show that the effect size in the first half of the academic year (r^2= 0.33) was larger than that (r^2= 0.13) in the second half, which points to a relatively more unstable L2 system in the very beginning stage of learning an additional language. To make the difference in degree of variability easier to understand, the changes in fluency scores of student A who had the lowest variability in fluency and student B with the highest were selected and visualized in Figure 1. We can see that while student A does not show much variation in the fluency scores, student B demonstrates quite ambitious peaks and dramatic drops.

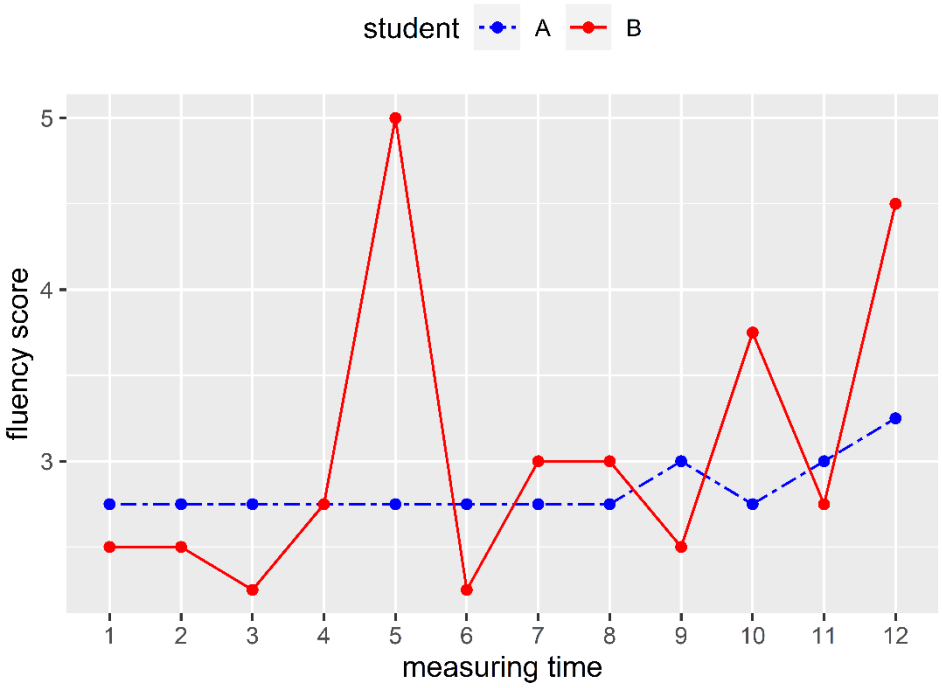

Figure 2.1: Low vs. high variability in fluency

This finding could be a result of the competition for limited resources, such as language learning motivation and time invested in L2 learning. Fluency in the current paper was measured based on text length. A text might be longer, for example, when the learner is able to access the L2 system more easily during that period of time, which in turn may be related to the amount of L2 exposure or use. While this could be a common phenomenon among all language learners, it might be more obvious among L2+L3 learners compared to L2 learners, because the former has one more language to cater to within the same amount of time available. Meanwhile, the motivation to write could also fluctuate from time to time, which may lead to a fluctuation on the length of the texts. As suggested by Henry (2010), the motivation to learn the two languages of the simultaneous L2 and L3 learners are not isolated nor evenly distributed. Simply put, an increasing L3 motivation during a certain period may lead to a decline in the L2 motivation and result in less time and energy spent on to L2 learning and use. From this point of view, variability is explained here as a 'symptom' of the temporary competition for limited resources.

However, variability may also be taken as a sign of progress. In another study within the same project, which set out to explore the best predictor among motivation, aptitude, working memory, or variability for the final L2 writing proficiency and gains, variability was found to be the best predictor (Huang, et al., 2020b). In other words, the more variable the learners were in their writing scores, the higher the scores in writing proficiency were. As in the current study, the

L2+L3 learners (M=3.31, SE=0.73) had significantly higher scores than the L2 learners (M=3.02, SE=0.49) on average fluency scores ($t(11)=6.08$, $p<0.001$). Thus, from this point of view, it is plausible to infer that less stable systems are more likely to develop/progress.

In sum, the results of the current study suggest that learning an additional L3 alongside the L2 does not hinder nor slow down the learning of the L2, but it does seem to lead to an initial destabilization of the L2 system, which appears to level off as time goes on. And this destabilization of the L2 system does not necessarily mean a negative unstable status of the system; instead, it may be a sign for the possibilities of development in the long run.

2.7 Conclusion and limitations

Language learning is a complex and dynamic process, and language learners are complex dynamic systems. Learning two foreign languages simultaneously makes the system and the learning process even more complicated. From a CDST view, the development of any complex dynamic system is dependent on limited resources and variability is viewed as a "potential indicator of the ongoing processes" (Van Geert & Van Dijk, 2002, p. 341). Based on this view, we hypothesized that due to the limitation in resources, the L2+L3 learners would improve to a lesser degree than the L2 learners. In addition, due to the competition in language learning resources, the L2+L3 learners would show different patterns of variability. Our findings showed that there were no differences between L2+L3 learners and L2 learners in holistic gain scores, not in the first year nor in the second year. These findings rejected our hypothesis that L2+L3 learners would gain less, and indicated that learning an L2 and L3 simultaneously does not hinder the development in L2. The variability analysis, showed that L2+L3 learners were more variable in L2 fluency scores, which suggests that learning an L3 makes the developing L2 system less stable, but it is not necessarily a negative sign, as other studies (Lowie and Verspoor 2019) have shown that more variable learners tend to gain more.

One of the major limitations of the current study is that it did not trace the dynamic changes on the individual learner factors, such as motivation, and the average time invested in each language outside the classroom. Such fluctuations might help to explain the greater instability in L2 fluency in the L2+L3 leaners. Future studies are encouraged to combine a dynamic observation of individual factors and writing development with variability analyses in order to explore how the individual factors might influence L2 learners and L2+L3 learners differently.

Another limitation of the study is that the results that L2+L3 learners did not gain less than the L2 learner in L2 writing cannot be generalized to other language tasks, for example, real-time tasks such as timed production or comprehension tasks. Those tasks involve real-time processing under time pressure and may be more demanding for cognitive abilities such as working memory span. Given the findings from another study in the current project that L2+L3 gained more than the L2 learners in working memory capacity, it would be an interesting topic for a future study to investigate whether L2+L3 learners differ from L2 in more real-time language processing tasks due to an enhancement in cognitive abilities such as working memory.