JYVÄSKYLÄ STUDIES IN HUMANITIES

314

Sanna Olkkonen

Second and Foreign Language Fluency from a Cognitive Perspective

Inefficiency and Control of Attention in Lexical Access

JYVÄSKYLÄN YLIOPISTO

JYVÄSKYLÄ STUDIES IN HUMANITIES 314

Sanna Olkkonen

Second and Foreign Language Fluency from a Cognitive Perspective

Inefficiency and Control of Attention in Lexical Access

Esitetään Jyväskylän yliopiston humanistis-yhteiskuntatieteellisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston vanhassa juhlasalissa S212
toukokuun 27. päivänä 2017 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Humanities and Social Sciences of the University of Jyväskylä,
in Auditorium S212, on May 27, 2017 at 12 o'clock noon.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2017

Second and Foreign Language Fluency from a Cognitive Perspective

Inefficiency and Control of Attention in Lexical Access

JYVÄSKYLÄ STUDIES IN HUMANITIES 314

Sanna Olkkonen

Second and Foreign Language Fluency from a Cognitive Perspective

Inefficiency and Control of Attention in Lexical Access

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2017

Editors
Lea Nieminen
Centre for Applied Language Studies, University of Jyväskylä
Pekka Olsbo, Sini Tuikka
Publishing Unit, University Library of Jyväskylä

Cover picture "Street art in Kraków" by Sanna Olkkonen (2011)
Permanent link to this publication: http://urn.fi/URN:ISBN:978-951-39-7062-8

URN:ISBN:978-951-39-7062-8
ISBN 978-951-39-7062-8 (PDF)
ISSN 1459-4331

ISBN 9978-951-39-7061-1 (nid.)
ISSN 1459-4323

Jyväskylä University Printing House, Jyväskylä 2017

ABSTRACT

Olkkonen, Sanna
Second and foreign language fluency from a cognitive perspective: Inefficiency and control of attention in lexical access
Jyväskylä: University of Jyväskylä, 2017, 72 p.
(Jyväskylä Studies in Humanities
ISSN 1459-4323; 314 (nid.) ISSN 1459-4331; 314 (PDF))
ISBN 978-951-39-7061-1 (print)
ISBN 978-951-39-7062-8 (PDF)
Diss.

The present dissertation focuses on the fluency of second and foreign language (L2). Second and foreign language use is often assessed by its fluency. However, the connections between language proficiency and fluency are far from clear. The fluency reasearch has tried to establish these connections, but results are varied. In the current thesis, this question is explored within the theoretical framework of Norman Segalowitz's division of three types of fluency: cognitive, utterance, and perceived fluency. The overall proposition here is that in relation to L2 proficiency, cognitive fluency provides important insights. This is based on the theories of limited cognitive resources, which state the more efficiently the lower-level processes work, the more resources will be available for the higher-level processes. Therefore, the fluency with which a language is used may in part depend on the resources available for both levels of processing.

The thesis comprises three journal articles and a synthesis of them. As a measure of cognitive fluency, lexical access tasks were used. Fluency was operationalised as the efficiency of lexical access (speed and accuracy) and the resources available for the control of attention. Speed and accuracy in both word recognition and word retrieval in L2 were measured. The participants were Finnish learners of English as a foreign language from Grades 4, 8, and 11; and a group of Russian-speaking learners of Finnish as a second language from primary school. Furthermore, the relationship of L2 lexical access to L1 lexical access was examined.

In Article I, the speed of lexical access was found to explain a substantial amount of variance (20–45%) in L2 reading comprehension and writing performance. Word recognition rates explained literacy skills in the younger readers, whereas word retrieval rates accounted for the literacy skills in the older students, indicating a difference in the skill profiles. In Article II, the accuracy of lexical access added 10–17% to the explained variance of the literacy skills, confirming the importance of accurate lexical access in addition to the its speed in literacy skills. Furthermore, fluent L1 lexical access was found to explain variance in L2 literacy skills across all grade levels. Article III examined inaccuracies in lexical access, which were categorised as resulting from either inefficiency or attention-control. These were shown to have different distributions depending on the level of language proficiency. The overall results thus highlight a componential view of fluency, with its distinct aspects connecting to L2 proficiency in more varied ways than have been considered thus far.

Keywords: second and foreign language, cognitive fluency, word recognition, word retrieval, literacy skills, error analysis

Author's address Sanna Olkkonen
Centre for Applied Language Studies
University of Jyväskylä, Finland
sanna.olkkonen@jyu.fi

Supervisors Lea Nieminen, PhD
Research coordinator
Centre for Applied Language Studies
University of Jyväskylä, Finland

Prof. Paavo Leppänen
Department of Psychology
University of Jyväskylä, Finland

Reviewers Prof. Judit Kormos
Department of Linguistics and English Language
University of Lancaster, UK

Ass. Prof. Patrick Snellings
Department of Psychology
University of Amsterdam, the Netherlands

Opponent Ass. Prof. Patrick Snellings
Department of Psychology
University of Amsterdam, the Netherlands

ACKNOWLEDGEMENTS

First of all, I am grateful to my supervisors Lea Nieminen and Paavo Leppänen. Lea always had encouraging words for the times when they were needed (which was often) and meticulous comments on the multitude of paperwork included in the process. Paavo's knowledge is impressive and his detailed questions helped me to refine my thoughts. I also have to offer my thanks to Ari Huhta for taking me in on the project, when all my plans were still very vague, as well as for acting as my first supervisor and helping me to design the inaccuracy categorisations. I want to express my thanks to the reviewers of this thesis, Professor Judit Kormos and Assistant Professor Patrick Snellings. Their expertise in their respective fields and their comments helped me to improve this work a great deal. Their input guided me to clarify many of the imprecise conceptions and overlooked issues in the manuscript.

My gratitude goes to my co-author and tireless advisor Kenneth Eklund, who always had time and patience to answer my silliest questions and listen to my ideas, no matter how far-flung. I am also deeply indebted to Rob Schoonen, for all the comments, discussions, and the time he devoted during my visits and inbetween. And overall, for making it possible for me to visit the inspiring athmosphere of University of Amsterdam and Radboud University Nijmegen, which offered me truly great times, both academically and personally.

I would like to warmly thank Pauliina Peltonen for the travel company and insipiring discussions: your ability to find sense in my ramblings is amazing. Being a distance student is somewhat lonely, but discussions with colleagues, on issues academic and non-academic alike, have meant a lot: thank you Tarmo Ahvenainen (e.g. Keira's acting talents), Teppo Jakonen (to Sohwi?), and Dmitri Leontjev (linguistic paper in Quenyan still pending). Tessa Spätgens, dankuwel voor weerberichten, nederlands les en gezelschap. Thank you Hanna Lantto, as who else would invite me to a party where I can dress up as my favourite linguistic phenomena? (Exclusive first person plural.) I am grateful to Paula Salmi for the comments and discussions when embarking on the slippery road of error analysis (your advice: Don't do that.). Thank you Jaana Haapasalo for your help in the coding of inaccuracies. For giving me a push to start in the first place, I wish to thank Eero Voutilainen and Mia Halonen, from the distant Kotus days.

For the financial support during my doctoral studies, I am grateful to the Faculty of Humanities and Centre for Applied Language Studies (CALS) of the University of Jyväskylä, and to the Ella and Georg Ehrnrooth Foundation. I also thank the DIALUKI project for employment in data gathering and the possibility to use the impressive amount of data my thesis draws from. I wish, further, to thank CALS, the Faculty of Humanities, Langnet National Doctoral Programme, and International Symposium on Bilingualism for travel grants - travelling to conferences must be one of the best things in doing research.

In the topsy-turvy world of academia, other sources of income often come in handy. Therefore, I wish to express my thanks to my bosses Merja Vannela and Sirpa Rönkkö for the flexible workhours during these years and lovely students. Furthermore, it would have been very difficult indeed to get anything finished without the office space provided by Mervi M. for which I am forever grateful; without that, only squeaky-clean cupboards would attest to my time as a graduate student.

Lastly, my dear friends, Iris, Kake, and Sanna M. To remain sane, you must have the possibility to discuss work, and then to totally forget about it. Olli: osta minulle juoma, laula minulle laulu, ota minut sellaisena kuin olen. My parents, as you have made me what I am.

Helsinki, April 2017
Sanna Olkkonen

LIST OF ORIGINAL PUBLICATIONS

This doctoral thesis is based on the author's three original publications, one of which consisted of a joint paper with two co-authors. The papers are referenced as follows:

Article I Olkkonen, S. (2013). Speed in cognitive tasks as an indicator of second/foreign language reading and writing skills. *Estonian Papers in Applied Linguistics*, 9, 195–208.

Article II Olkkonen, S., Eklund, K. & Leppänen, P. H. T. (manuscript). Fluency of lexical access in L1 and L2 in relation to L2 literacy.

Article III Olkkonen, S. (2017). Processing limitations in L2 fluency: Analysis of inaccuracies in lexical access. *APPLES: Journal of Applied Language Studies*, 11, 19–41.

FIGURE

TABLE

CONTENTS

1 INTRODUCTION

It is increasingly rare nowadays to find a person who can go through life without having to interact in at least one additional language. Worldwide, multilingualism is more the rule than the exception. The fluency with which we use these languages, whether when speaking, listening, writing, or reading, affects how we are perceived by others, one way or another. Most of us are familiar with the annoying feelings of incompetence and frustrating childlikeness when trying to express ourselves as capable adults with a language in which we are not fluent. Listeners often seem to make judgments about our intelligence and the reliability of our message based on their initial impressions of how effortlessly we are able to communicate our intentions. Sometimes native speakers express their lack of patience towards people with even quite proficient foreign language skills, for not being used to 'broken language'. This is especially prevalent in the language areas with yet little experience on second language users, as is the case in Finland. However, the concept of fluency is a moving, slippery thing. It seems easy to notice when it is not present, but the research aimed at pinning down the exact features of which it consists has not reached a consensus. Fluency is often used as a factor when assessing proficiency in foreign language classrooms, but a key question remains: can the fluency with which we use a foreign language really inform others (researchers included) about our abilities and language proficiency?

1.1 Fluency in second and foreign language

One possibility for studying fluency is not to treat it as a solid construct, but rather to examine it from a componential viewpoint, as comprising different dimensions of performance. There is the type of fluency that we can observe when someone uses a language, whether mother tongue (L1) or second or foreign language (L2), and the resulting opinions we form based on that observation. There is also the measurable fluency of pauses, rate, and errors, both in speech and in writing. Furthermore, there exists the fluency of mental operations responsible for language production, understanding, and integration of information. This kind of division is proposed by Segalowitz (2010) in his seminal book. He defines the connections of these different aspects of fluency as follows:

> What *features of L2 oral performance* serve as reliable indicators of how efficiently the speaker is able to mobilize and temporally integrate, in a nearly simultaneous way, *the underlying processes of planning and assembling* an utterance in order to perform *a communicatively acceptable speech act*? [italics added]. (2010: 47)

He cautions that one should delineate these three aspects of fluency: they interact but are not synonymous. For example, the fluency of underlying cognitive operations is not restricted to speaking or even to language use, as it also indicates the ease and speed of reading and writing, motor skills, perception, etc. (see also Schmidt, 1992). If we want to examine which fluency features are the most informative of L2 proficiency, we must first determine how and what to measure. Information processing theories suggest that any skilled performance builds on co-operation of automatic and controlled processes (see Schmidt, 1992; Segalowitz, 2000, 2003, 2010). Cognitive fluency may be seen as a reflection of the balance of these two, and therefore, by measuring these we may be able to tap into the efficiency of cognitive operations. Automatic processes are quick and accurate, but difficult to control; controlled processes, on the other hand, are conscious and flexible devices for monitoring, maintaining goals, and making decisions that ascertain successful communication in real-life environments (Schmidt, 1992; Segalowitz, 2000). Consideration of these aspects may offer valuable insight into the measurement of cognitive fluency.

A common approach taken for measuring any level of language use is an oral task assessment (although this is by no means the only way; see e.g. Manchón, Murphy & Roca de Larios, 2007). There are, however, certain caveats to this approach. One question concerns the extent to which oral tasks are able to tap into L2 proficiency, or whether oral fluency is merely "in the ears of the beholder", as Freed (2000) submits. We must, therefore, determine a way to measure fluency that does not solely depend on listeners' impressions. A method to control for this would be to demonstrate that the fluency measures used are connected to wider language skills. Interestingly, and slightly counterintuitively, certain oral reading tasks (e.g. reading words aloud) have been shown to be good indicators of language proficiency (e.g. reading comprehension) (Nation & Snowling, 2004; see also, Verhoeven, 2000), more so than silent reading tasks (e.g. L. S. Fuchs, Fuchs, Eaton & Hamlett, 2000, unpublished data cited in Fuchs, Fuchs, Hosp & Jenkins, 2001). This suggests that the interaction does not stem solely from the ability to read; a possible mediator may be the fluency of underlying cognitive processes required in both tasks. From L1 reading studies, we know that word retrieval plays an important role in reading development (see e.g. Bowey, Storey & Ferguson, 2004; Di Filippo et al., 2005; Georgiou, Parrila, Cui & Papadopoulos, 2013; Heikkilä, Närhi, Aro & Ahonen, 2009; Kirby, Parrila & Pfeiffer, 2003; Protopapas, Fakou, Drakopoulou, Skaloumbakas & Mouzaki, 2013; Puolakanaho et al., 2007). To tap into the connections between fluency and proficiency, therefore, lexical access (as both word recognition and retrieval) may serve as a good tool for measurement.

A second question that arises is that if oral tasks are used to measure language use, which features of fluency are the most informative concerning proficiency? There seems to be an interesting discrepancy in play in that both naïve and expert listeners are in fact good at determining the proficiency level of L2 learners — but it has proven difficult for researchers and educators to pin down the objective criteria. For example, in the Common European Framework of Reference for Languages (CEFR, Council of Europe, 2007), the guidelines for assessing fluent speech are the lack of pausing and "natural, smooth flow of language" (p. 28; cf. Koponen & Riggenbach, 2000). These formulations are problematic, as they are prone to be subjective and under-defined. What exactly is natural flow of language? As Lennon (1990) has pointed out, L1 speech is not without interruptions either, and includes pausing, self-corrections, and repetitions. If fluency is used in assessing proficiency, the measurements should not be dependent on individual speaking styles or conversational skills. To eliminate these effects, the measures should be strictly formulated, and it should be ensured that they tap into cognitive operations rather than speaking skills. Lexical access measurement may serve as a possibility in this respect as well. Furthermore, controlling for L1 behaviour is required.

1.2 Outline of the thesis

The answers to the aforementioned questions are sought in this thesis, which comprises results from three research articles and a discussion on the whole which they form. In these articles, I have examined different sides of cognitive fluency: automaticity or **efficiency** (speed and accuracy) and **attentional processes** in L2 lexical access (both word recognition and retrieval) in second and foreign language learner groups consisting of four different levels. In Article I, I concentrated on the **speed of lexical access** and its relationship to **L2 literacy** (L2 reading comprehension and writing). In the scope of this thesis, L2 is considered to cover all the additional languages (second, third, foreign) that are learnt after the first one (L1), when not explicitly stated otherwise. Second, in Article II, I examined the role of **accuracy in lexical access**. When listening to what the participants actually produced when recognising or retrieving words, it was noticeable that even if someone was fast in performing the tasks, they may have been in fact producing something quite different from the actual items. A review of previous literature revealed that accuracy had rarely been looked into in relation to lexical access (cf. Bowey et al., 2004; Di Filippo et al., 2005; Dufva & Voeten, 1999). In Article III, I analysed the **inaccuracies in lexical access** and categorised them in relation either to lack of efficiency, or to attentional processes. Furthermore, similar lexical access tasks were conducted in the L1 to measure the connections of **L1 fluency** to L2 fluency (Articles I–III) and L2 literacy (Articles I–II).

To summarise, the research questions examined in the articles are as follows:

1. How much does the speed of both L1 and L2 lexical access explain the variance in L2 literacy skills of Finnish learners of English, and Russian learners of Finnish?
2. Does the accuracy of both L1 and L2 lexical access uniquely explain variance in the L2 literacy skills of Finnish learners of English, after accounting for the speed of lexical access?
3. Does L2 proficiency affect the amount and type of inaccuracies in lexical access of Finnish learners of English?

Methodologically, my aim has been to combine the theoretical frameworks of **applied linguistics** and **cognitive psychology**. The advantage of this approach lies in the opportunity to use the standardised and validated materials provided by the psychological research and combine them with a more qualitative linguistic approach. Fluency is such a transient phenomenon that this combined methodology was considered to reveal a fuller picture of how fluency in L2 should and could be studied. Therefore, a mixed-methods approach was applied. The statistical analyses helped to examine the differences and the patterns of language use at the different proficiency levels in a general way (Articles I–II). On the other hand, the more qualitative analyses (Article III) seek to reveal more individual variation and interesting phenomena that, although mostly not frequent enough to show significant effects in the statistical analyses, still offer valuable information for further research. The data was a part of the larger DIALUKI project (see **Chapter 3.1**). In Articles I and III, I was the sole author. Article II is a joint paper; I was responsible for the original idea and data coding, and was also in charge of the writing of the paper. The research questions, analyses, and interpretation of the results were performed in dialogue among all the authors. In the last paper, Article III, the inaccuracies in two lexical access tasks were categorised using a design I developed in collaboration with Prof. Ari Huhta (explained in **Chapter 3.3.2**).

The organisation of the remainder of the thesis is as follows. First, in **Chapter 2**, I cover the relevant theoretical discussion concerning fluency, as well as lexical access and its relationship to literacy skills. I look briefly at interactions between L1 and L2 fluency. In **Chapter 3**, I introduce the design of the present study, including data gathering, my personal contribution, the participants, and the methods. In **Chapter 4**, I present the main results from Articles I–III. **Chapter 5** offers a discussion of the implications of these results, and suggestions for further research. The original papers are found at the end of the thesis as appendices.

2 THEORETICAL BACKGROUND

In this section, the main theoretical frameworks behind the thesis are discussed. First, I examine the theoretical foundations of fluency, and concentrate on the cognitive perspective in connection with limited processing capacity in language use. Then, the impact this viewpoint has on the definition and the measurement of L2 fluency is explored. This is followed by the theoretical model used for fluency in this thesis. The reasoning for using lexical access as a measurement tool for cognitive fluency, as well as its relationship with literacy skills are discussed next. Lastly, the connections between L1 and L2 fluency are considered briefly.

2.1 Fluency in language use

The most important question in my thesis is what can be considered as *fluent* second language use? The results from the surface measurements (**utterance fluency** per Segalowitz's terminology) vary, not least because of the wide array of the individual phenomena studied. Utterance fluency has often been studied within the framework of CAF (complexity, accuracy, and fluency; see Housen & Kuiken, 2009), and most researchers base their operationalisations on Skehan's (2003) model of repairs, breakdowns (pauses), and speed. Pausing and speed have revealed the most robust results and are considered to be good markers of (dis)fluency (Bosker, Pinget, Quené, Sanders & De Jong, 2013; Freed, 2000; Kahng, 2014). However, even for these markers of disfluency, caution should be taken: some phenomena traditionally measured and associated with disfluent speech have been shown to correlate more to individual speech styles and strategies (e.g. Walczyk, Marsiglia, Johns & Bryan, 2004). De Jong, Groenhout, Schoonen, and Hulstijn (2015) showed that it is not necessarily the length of pauses but the amount of them (both filled and silent) that is most reliably related to proficiency. Furthermore, the relationship between repairs (false starts, repetitions, replacements, and reformulations) and L2 proficiency has been questioned on many occasions (Engelhardt, Corley, Nigg & Ferreira, 2010; Gilabert, 2007; Kormos, 1999; Lennon, 1990). These results highlight the need to determine which of the surface fluency features are the most reliable reflections of processes underlying L2 proficiency.

When one assesses fluency, it is virtually impossible to avoid encountering **perceived fluency**, that is, the interpretations which the listeners make on the speaker's abilities based on their surface fluency (Segalowitz, 2010). Again, not all pauses and hesitations are considered as disruptions by even untrained raters, as, for example, Bosker and colleagues (2013) reported that fluency ratings were mostly affected by the amount and length of silent pauses, following a pattern similar to that of the expert studies. Interestingly, they also found that even the untrained raters, although sensitive to repairs, did not allow these to influence their ratings markedly. Perceived fluency is, nevertheless, only indirectly linked to cognitive processes via utterance fluency.

2.1.1 Cognitive fluency and its aspects

As Lennon (2000: 27) states, "For most speakers in most situations, processing demands, rather than deficient knowledge, will limit fluency". In their influential theory of information processing, Just and Carpenter (1992) describe that human cognition is limited, and these limits are set by working memory capacity. **Working memory** is a cognitive system responsible for processing and storing incoming information, and its capacity predicts achievement in many higher-order tasks (Baddeley, 1996). Because of the limited capacity of working memory, the more fluently the underlying processes operate, the more capacity there is available for the higher mental processes, such as monitoring and comprehension (Grabe, 2009; Kirby et al., 2003; LaBerge & Samuels, 1974; Perfetti, 1985). **Cognitive fluency** may serve as one indication of this processing capacity, as it is defined as "the efficiency of operation of the underlying processes" of language use (Segalowitz, 2010: 165; see also, Schmidt, 1992). Fluency, in Schmidt's (1992: 366) words, "can validly be described as the control of mostly automatic processes by selective attention". Therefore, any fluent cognitive performance requires a balance between automatic and controlled processes (Segalowitz, 2000, 2010: 91).

In language use, **automaticity** of sub-processes entails different prerequisites for reading, listening, and speaking. In reading, for example, the words in the text need to be recognised efficiently, that is, fast and accurately (Ehri, 1991; Segalowitz, 2010). In listening, the links between the grammatical categories and the conceptual interpretations should work automatically (Segalowitz & Frenkiel-Fishman, 2005). In speaking, automaticity of word retrieval is required to be able to attain a 'normal speech rate' (see e.g. Schmidt, 1992). Efficiency and automaticity are often used synonymously, but Segalowitz (2000, 2010) has stressed that automaticity requires more than just an acceleration of the operation; there should be mental restructuring as well (although for cautioning remarks, see Hulstijn, Van Gelderen & Schoonen, 2009). Therefore, the efficiency of the mental processing can be seen in the speed, accuracy, and stability with which these operations are performed. Development of efficiency with repetition and training could be interpreted as a process, the goal of which being automaticity. Automaticity of a process means that it works unconsciously, unintentionally, involuntarily, and effortlessly. It is thus also more immune to interference from outside sources (Grabe, 2009; Segalowitz, 2003). As efficiency develops only gradually, there may be, at first, a trade-off between the speed and accuracy of a process (shown in the variability of the responses). Only with practice may automatization develop, which in turn frees up cognitive resources for the more controlled, attention-requiring actions (e.g. Schmidt, 1992).

These **attentional processes** include, for example, focusing on and shifting between tasks, and there are several ways of conceptualising them (see Segalowitz, 2010: 93). In cognitive psychology, they fall under the domain of the central executive of working memory, which is responsible for directing attention, suppressing irrelevant information, monitoring, and combining and manipulating new information (Baddeley, 1996; Grabe, 2009: 35). The central executive concept is, furthermore, proposed to be linked to domain-general executive functions studied in the neuropsychological tradition (e.g. Baddeley, 1996; Diamond, 2013; Kieffer, Vukovic & Perry, 2013; Miyake et al., 2000). McCabe, Roediger, McDaniel, Balota, and Hambrick (2010) showed that the tasks used to measure working memory capacity and executive functions correlate highly (r = .97); in fact, they can be argued to measure a common underlying ability. The main components in both executive control theories and executive functions are attention and control thereof. Executive functions are generally divided into three sub-processes: **inhibition**, **monitoring,** and **shifting** (Diamond, 2013; Miyake et al., 2000). This categorisation offers a possible, well-structured framework for studying the effect of attentional processes in language use as well. In this framework, to be a fluent language user means that one must be able to efficiently inhibit irrelevant information, to shift between different requirements of the situation, and to monitor one's performance. The studies concerning the role of executive functions in L2 learning are hardly abundant (but see Bialystok, Craik & Luk, 2008; Marian, Blumenfeld, Mizrahi, Kania & Cordes, 2013). This is notwithstanding Segalowitz and Frenkiel-Fishman's (2005) suggestion that the ability to control attention is one of the most important mechanisms that differentiate expert language users from beginners. My proposition is that the framework of executive functions may help in interpreting some seemingly conflicting fluency phenomena.

2.1.2 Fluency model of the current thesis

Based on the theoretical considerations concerning the features of fluency, I propose the theoretical construction of the concept of fluency that follows (see **Figure 1.**). The construct is based on Segalowitz's framework (2000, 2010), with few inclusions. Presented therein are suggested interconnections among the different aspects of fluency. The limited processing capacity (Just & Carpenter,

1992) is the overall mechanism that determines the limits for fluency. Fluency itself is divided into cognitive fluency, utterance fluency, and perceived fluency (Segalowitz, 2010). The cognitive fluency is apparent in the phenomena connected with efficient processing (automatization or the lack thereof) and attentional processes (Segalowitz, 2000). The efficiency component is proposed to be measurable by speed, accuracy, and stability of production. In the current study, the stability measurements were not be included, as the reaction times across items were not counted (see Segalowitz, 2010). As for the attentional processes, my proposition is to operationalise these along the framework of executive functions, that is, a division into monitoring, inhibition, and shifting (Miyake et al., 2000). For utterance fluency, Skehan's (2003) framework is presented here, but this is not the object of the current study. However, two features usually included in utterance fluency phenomena (repairs), self-corrections and repetitions, are included: they are hypothesised to be more strongly related to monitoring ability (attentional process) than to surface speaking skills. Since this thesis concentrates on cognitive fluency, the following chapters consider the measurement of this fluency, as well as the connections to wider language skills and between L1 and L2 fluency.

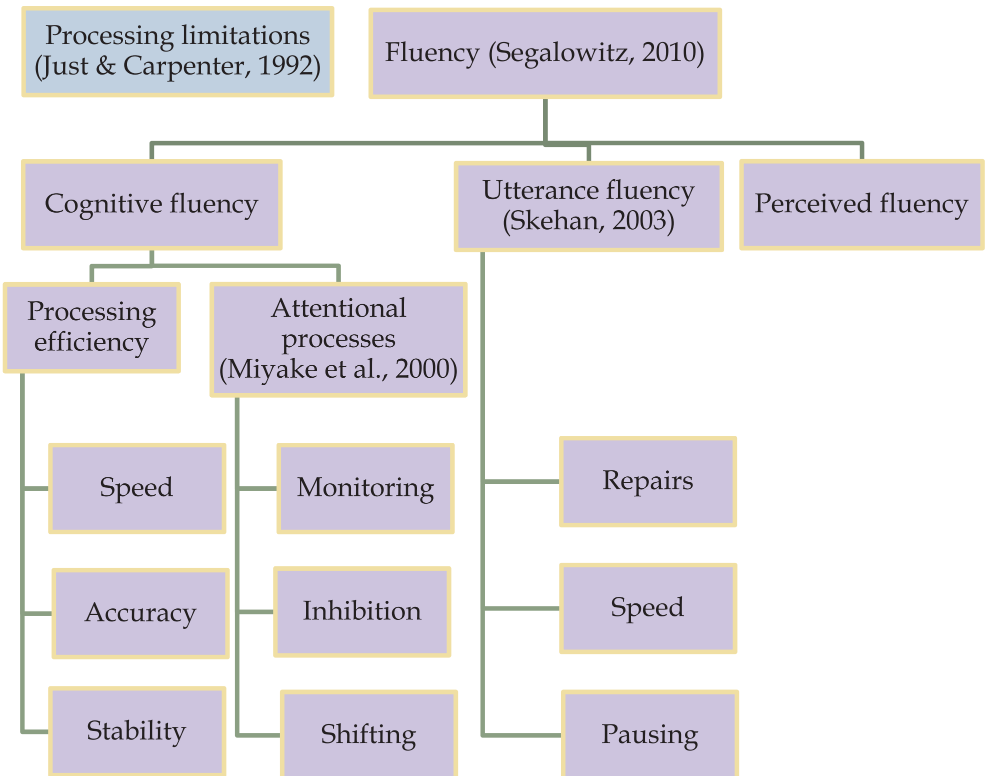

FIGURE 1 The construction of the fluency concept applied in the thesis.

2.1.3 Inaccuracies as reflections of processing capacity

Engelhardt et al. (2010) suggest that the impact of processing capacity may best be researched in the instances when it is lacking, that is, in the form of errors. From the fluency model introduced above, it therefore follows that if cognitive fluency is limited by the requirements of both efficient language skills and control of attention, then the way to tap into cognitive fluency may be found in examining the breakdowns of these separately. Those breakdowns that are caused by limited proficiency should be distinguished from those that are more related to limitations in the capacity of cognitive operations (see also Kormos, 2000). However, the term 'error' has been used in accuracy-oriented literature (Di Filippo et al., 2005; Ehri, 1991; Gilabert, 2007; Kormos, 2000), implying a dichotomous distinction between a correct and an incorrect answer, which does not seem to cover sufficiently the conscious control of attention. The term 'disfluency', on the other hand, is used in utterance fluency research to refer to surface properties of fluency, and it is often considered to be separate from accuracy (e.g. Bosker et al., 2013; Kahng, 2014). Disfluency is also used by speech pathologists to refer to stuttering, for example. Hence, in the current thesis, the chosen term is **'inaccuracy'**, as it seems to cover the use intended in the thesis in the most appropriate sense. It is chosen both to avoid confusion with the operationalisations of other theories and, further, to highlight that the current categorisation includes inaccuracies related to different sources: to both inefficiency and attentional processes. This is an important distinction, as the focus of the present study is on the limitations of the capacity of cognitive operations rather than on those of proficiency.

The **efficiency** of language use can be measured by speed and accuracy of the performance. If language skills are limited, and the relevant sub-processes are not yet automatized, there may be inaccuracies resulting from either of these. Therefore, there may be some competition, that is, a trade-off between aspiring for speed and aspiring for accuracy (Grabe, 2009: 292; Kame'enui & Simmons, 2001; Walczyk et al., 2004). There are some tentative findings on individual fluency profiles (Gernsbacher, 1993), and on differences between learners who favour either speed or accuracy in their performance (Catts, Gillispie, Leonard, Kail & Miller, 2002; Jeon, 2012; Salmi, 2008). If a potentially automatic process, such as lexical access, is not yet automatized, it requires cognitive resources, and less capacity would seem to be available for the executive processes, such as monitoring the appropriateness of speech (Kahng, 2014; Kormos, 2000; Levelt et al., 1999). Only with sufficient language proficiency can the higher-level cognitive skills (such as monitoring or the more strategic reading skills) be utilised (cf. Walczyk et al., 2004).

The problems in **attentional processes** relate to the inability to manage or direct attention (Segalowitz, 2000, 2010: 92). Jarvis et al. (2014) studied inhibition and shifting in L2 learners' data and found that both phenomena emerged only at later stages in language learning. The relationship of these with language proficiency also strengthened over time and increasing proficiency. In an experiment that Segalowitz and Frenkiel-Fishman (2005) conducted, they found that the speed of L1 and L2 attention-control in a task-switching experiment accounted for 59% of the variance in L2 proficiency. However, interestingly, there is also a possibility that control of attention might in fact lead to inaccuracies when there are more resources to use it. Thus, it is hypothesised that there are types of errors that are connected with the development of language skills (Ehri, 1991; see also Engelhardt et al., 2010; Protopapas et al., 2013; Schmidt, 1992). For example, Kormos (2000) reports differences in the type of errors between L1 and L2 speech.

In terms of the attentional processes, monitoring one's own speech is difficult with low proficiency, as the resources available are required for more basic processes such as word retrieval (e.g. Kormos, 2000; Segalowitz, 2010). Therefore, it is probable that inaccuracies such as self-corrections are not present in the production of beginning language learners. For shifting between the tasks, Segalowitz and Frenkiel-Fishman (2005) report that there is more shift-cost when switching from L2 to L1 than vice versa. This is interpreted as reflecting the need to suppress L1 more to be able to perform in the less automatized L2 (cf. Meuter & Allport, 1999). Shifting between the activated categories may thus be more difficult when the activation of a lexical category is stronger. In the description of their automaticity model, Samuels and Flor (1997) state as well that a drawback of automatization of a process is the greater difficulty of suppressing and controlling it. This can be illustrated with the Stroop effect. In the Stroop task (Stroop, 1935), a colour word is printed in a congruent ink (the word 'RED' is printed in red ink) or an incongruent ink (the word 'RED' in blue ink). The participant is to name the colour of the ink, and it takes longer and is more prone to errors when the colour of the ink is incongruent with the word (for a review, see MacLeod, 1991). Marian et al. (2013) studied this effect on participants' three languages, and found a stronger Stroop effect for L1 and L2 than for L3 (the least proficient language). The researchers interpreted this as signifying that the activation of the weaker language was easier to inhibit than that of the stronger languages (see also Gernsbacher, 1993). Another example of inhibition difficulties with automaticity is available from eye-tracking studies. Reading utilises a parafoveal preview, a pre-processing of words just outside of the fixation point, which reduces the time needed to process that word and ensures a fluent reading process (Pollatsek, Rayner & Balota, 1986). This may also influence the processing of the fixated word (Simola, Holmqvist & Lindgren, 2009; also in listening: Ellis, 2002). Jones et al. (2008) found that, even in a naming speed task, there is sensitivity to some visual and phonological information from the material in a parafoveal preview. As the lexical access processes are to a great extent automatic, at least in L1, this may lead to some pre-processed information influencing the production of the fixated item. This particularly occurs when there is a requirement for speeded performance. The result may be that the more advanced language users produce inaccuracies rather than the beginners.

2.2 Lexical access

As the framework for measuring cognitive fluency has now been established, the next question becomes, what kind of tasks could be used to measure it? Measurements of cognitive fluency are rare, and the manner in which it could be tapped into is still unclear. Kahng (2014), for example, has used a stimulated recall setting, where after a speaking task, the participants are asked to recall what they were thinking during the pauses and repairs of their speech. Recall is proposed to reflect cognitive processes, but is nevertheless a highly indirect measure. Segalowitz (2010: 75–6) proposes that fluency of **lexical access** might be an informative proxy for cognitive fluency; moreover, it is easy to measure. To measure lexical access, he has employed, among others, animacy judgment tasks, where the participants decide whether the nouns they see are animate or inanimate (Segalowitz & Frenkiel-Fishman, 2005). Theoretical support for using lexical access as an index of cognitive fluency may be found from the theories of word retrieval (Levelt, Roelofs & Meyer, 1999; more in **Chapter 2.2.1**), where accessing the lexicon is one point in language use in which the L2 experience differs quite radically from the mostly automatized L1 (see also Kahng, 2014). Research has revealed that, even though often measured by oral tasks, the efficiency of lexical access seems to serve as a reflection of mental processing (Denckla & Cutting, 1999; Di Filippo et al., 2005). It then follows that fluency of lexical access could be considered an index of cognitive fluency rather than a surface-level (oral proficiency) phenomenon. Compared to the longer monologue or dialogue settings usually used in utterance fluency research, lexical access may reveal more about efficient cognitive processing when the contextual and communicative aspects are reduced.

The term 'lexical access' itself has been used in varied ways, referring sometimes to word retrieval (e.g. Gholamain & Geva, 1999; Kirby et al., 2003; Manchón et al., 2007), and sometimes to word recognition (e.g. Akamatsu, 2008; Grabe, 2009; McCutchen, 2001; Snellings, Van Gelderen & De Glopper, 2002). The relationship between word recognition and retrieval is not clear either. Word retrieval is often studied with naming speed (see e.g. Bowey et al., 2004; Di Filippo et al., 2005; Georgiou et al., 2013; Heikkilä et al., 2009; Protopapas et al., 2013), and its connections to both lexical access and word recognition are much discussed. Numerous studies on naming speed have stressed that it is decisive in the development of word recognition: slow naming speed leads to slow word recognition (e.g. Di Filippo et al., 2005; Kail & Hall, 1994; Kirby et al., 2003). This is proposed to be caused by word-processing issues, which complicate the building of reliable orthographic representations of words (Torppa, Georgiou, Salmi, Eklund & Lyytinen, 2011). However, proportions reflecting the extent to which naming speed can explain word recognition vary greatly among researchers, from none (Catts et al., 2002) to up to 76% of text reading speed (Georgiou, Papadopoulous & Kaizer, 2014; see also Denckla & Cutting, 1999; Gholamain & Geva, 1999). Word recognition is described in Van

Gelderen et al. (2003) as a prerequisite for lexical access, and Närhi et al. (2005) and Logan, Schatschneider, and Wagner (2011) consider lexical access as a component of naming speed.

These considerations and differing views raise the question of whether word recognition and word retrieval are in fact, at least partly, reflections of different cognitive processes; or whether they might be both interpreted best as tapping into lexical access, but from different directions. Snellings et al. (2002) discuss the notion that word retrieval and word recognition can be considered to be related processes, in the way that both require efficiently operating connections to the mental lexicon and rely on retrieving the phonological form of words. Only the direction of the sub-processes is different (see also Gollan et al., 2011). They could therefore be examined as measuring different aspects of lexical access; this is the approach applied in the current thesis. The term lexical access is used when referring to these processes concurrently, as opposed to occasions when their particular differences are under focus. Next, I present what is known about word recognition and retrieval, and why these might connect to wider language skills, especially to literacy skills.

2.2.1 Theoretical background of word recognition and word retrieval

Word recognition entails mapping an orthographic form of a word in a text, or a phonological form in speech, with a lexical concept or representation (Balota, Yap & Cortese, 2006; Koda, 1996). There are two influential word recognition models to date. In dual-route processing (Coltheart, Rastle, Perry, Langdon & Ziegler, 2001), word recognition proceeds from print via two alternate processing ways. The first is the lexical route to semantic content, through which irregular words are reached; and the second is the route to phonological decoding, which is responsible for the reading of new words or pseudowords. The connectionist model (Seidenberg & McClelland, 1989), on the other hand, is based on an associative network of phonemes, letters, and morphemes. These connections are enforced by training, that is, automatization of the connections. The definitive answer for choosing between these two approaches is still lacking, but they both seem to be suitable for explaining distinct phenomena in word recognition (Balota et al., 2006). A common element in both theories, moreover, is the reliance on automaticity of word recognition (cf. Nation & Snowling, 2004).

In line with the dual-route model, Ehri's (1991) conceptualisation of fluent reading requires the recognition of most words in a text by sight, which means recognising them as wholes and not phoneme-by-phoneme. This sight vocabulary is proposed to develop with the phonological decoding of words, which, after being rehearsed, establishes robust access routes to memory, and the lexical access does not require phonological coding anymore. Words are thereby retrieved from memory quickly and accurately, and merely seeing them activates the phonological and semantic information directly. A fluent reader should, furthermore, be able to adjust the reading process and use both routes alternately and efficiently, recognising familiar words by sight and decoding new words by syllable or letters. In L1, word recognition works involuntarily and unavoidably, which has been demonstrated, for example, with the Stroop task (see **Chapter 2.1.3**).

In the case of L2 word recognition, automaticity develops more slowly, and several different phenomena have been suggested as influencing this development, such as orthographic distance and the positive and negative transfer between languages (see Koda, 1996, for a review). Typological differences may impact the reading process: inflectional languages, such as Finnish or Arabic, require more decoding skills, but to what extent the inflected forms are read morpheme-by-morpheme or stored as wholes is still open to debate (cf. Geva, Wade-Woolley & Shany, 1997). On the other hand, in the more analytic languages like English, whole-word reading strategies can and should be applied to a greater extent. Orthographic differences favour different reading strategies as well. The deep orthography of English does not phonetically reveal how to pronounce words, but in Finnish, the pronunciations are highly systematic and predictable (Aro & Wimmer, 2003). These differences affect which process is applied and to what extent in reading.

Word retrieval entails the production aspect of lexical access: mapping between a familiar concept or object and an existing lexical representation (Logan et al., 2011; Wolf, 1986). For word retrieval in speaking, the most widely cited framework is that of Levelt, Roelofs, and Meyer (1999). Word retrieval starts with a concept or idea the speaker wants to express, for which s/he then searches for a lexical entry in the mental lexicon. This is then mapped to a morphological form for which a phonological form is selected. The last phase is the articulation of the word. L2 models are built on these L1 models, and they highlight the points where the L2 experience diverts from L1, such as the question of whether L2 words connect to concepts directly or via L1 translations (Kroll & Sunderman, 2003; Segalowitz, 2010). For writing, there is no comparable model for word retrieval, although in Hayes's influential processing model of writing (1996), fluent lexical access is considered to be an integral component at all stages of writing (see also Manchón et al., 2007).

2.2.2 Fluent lexical access in reading

Is fluency of the lower-level processes then reflected in the wider language proficiency, as the theoretical questions in the first chapter asked? In the case of reading comprehension, the current view is a componential one, in that the efficiency of the lower-order components, such as lexical access, is a prerequisite for the higher-order ones, such as comprehension (see **Chapter 2.1.1**). This model relies heavily on Perfetti's (1985, 2007) reading model, which introduces efficient word recognition as a central component in reading ability. The model also emphasises the automaticity of well-developed connections. Therefore, it is not surprising that one of the best predictors of subsequent **L1 reading** difficulties has been shown to be problems in word retrieval (e.g. Puolakanaho et al., 2007). There are different kinds of word retrieval tasks (such as typing the first letter of a cue word, or naming words beginning with a certain letter), but the most widely used in L1 studies are the naming speed tasks, namely Rapid Automatized Naming (RAN: Denckla & Rudel, 1976) and Rapid Alternating Stimulus (RAS: Wolf, 1986). In the RAN tasks, the participants name aloud highly familiar letters, numbers, colours, or objects, presented in serial format, as quickly and accurately as they can. In the RAS task, these categories are mixed in such a way that the participant must alternate between the categories.

There exists extensive literature on the nature of the connection between naming speed and reading, but why exactly this connection is so prevailing is still much discussed. The studies have, among other things, shown that the connection between naming speed and reading skills is not explained by IQ (e.g. Catts et al., 2002), articulation speed (Denckla & Cutting, 1999; Di Filippo et al., 2005), or phonological processing (Cronin & Carver, 1998; Jones, Obregón, Kelly & Branigan, 2008). On the other hand, oral performance is proposed to be one important component in the relationship (Georgiou et al., 2013; for similar results in word recognition, see Fuchs et al., 2001). The common element between naming speed and reading has also been proposed to be a more general processing speed (Catts et al., 2002; Kail et al., 1999; Närhi et al., 2005), which would be in accord with the fluency approach proposed in **Chapter 2.1.3**. Thus, naming speed would be more a *reflection of* rather than *reason for* the processes underlying reading achievement. Kirby and colleagues (2003: 462) propose that naming speed may be "an index of how much and how quickly information may be entered in working memory" from long-term memory. Naming speed would, in this view, be reflected in reading skills because the more rapidly words are retrieved, the more time there is for decoding and integrating information before it is lost from working memory (e.g. Kail & Hall, 1994). It would tap, therefore, into the automaticity component of fluency (cf. Catts et al., 2002; Cronin & Carver, 1998).

On the other hand, the connection to reading seems to depend in part on the format of the task. The serial formats of RAN and RAS, as compared to presenting the same items discretely, have revealed consistent differences between good and poor readers (De Jong, 2011; Denckla & Cutting, 1999; Georgiou et al., 2014; Jones et al., 2008; Logan et al., 2011). These results are in line with Wolf's (1986) proposition that the connection of naming speed and reading might stem from sequential processing, as well as inhibition of the context and control of attention, which are required in both. Consequently, as seen in **Chapter 2.1.1**, these processes are examples of attention-control. Along similar lines, Denckla and Cutting (1999: 38) have discussed the possibility of naming tasks falling into the "language/executive (functions) crossroads". For example, Närhi and colleagues (2005) found that an executive function task (verbal fluency) explained small variance in naming speed. Gernsbacher (1993) has proposed that all comprehension in reading, in listening, and from pictures, relies on common cognitive mechanisms, one of which could be inhibiting irrelevant material. She also demonstrated that less-skilled L1 readers were less able to suppress inappropriate information. Kaakinen, Hyönä, and Keenan

(2002), on the other hand, showed that good readers were better able to concentrate on the relevant parts of the text while skimming, which means they have better control of their attention (see also e.g. McVay & Kane, 2012). The speed in the serial naming speed task has been found to increase more than in the naming of discrete items, and this is suggested to reflect how beginners process items one at a time; the skill to activate and deactivate (i.e. to suppress) material only slowly develops (De Jong, 2011; Logan et al., 2011; cf. Geva et al., 1997). Kieffer and colleagues (2013) examined the role of two executive functions, inhibition and shifting, in relation to lexical access and reading comprehension. Inhibition was measured with a Wisconsin Card Sorting Test, where the participants adjust their sorting of cards according to changing rules. For the shifting measurement, a numerical Stroop task was used. Both contributed uniquely to the outcome measures. Therefore, the performance in the naming speed tasks seems to reflect the inner workings of attention-control mechanisms; thus it seems when naming, also aspects of attention-control may be examined.

Despite having a strong relationship with L1 reading, word retrieval tasks have been rarely used in **L2 reading** research. A notable exception are Geva and her colleagues (cf. Geva, Wade-Woolley & Shany, 1997; Geva & Massey-Garrison, 2013; Nassaji & Geva, 1999). Gholamain and Geva (1999) measured naming speed in both L1 (English) and L2 (Persian) in beginning readers. They found that naming speed explained word recognition within the languages and, further, that L1 naming speed explained a small amount of variation in L2 word recognition. Individual differences in the speed of lexical access accounted for the word reading results, regardless of their familiarity, thus reflecting the efficiency of lexical access. They did not, however, compare naming speed to more general reading skills. On the other hand, a plethora of word recognition tasks have been used in relation to L2 reading comprehension (for an overview, see Koda, 1996; also Jeon & Yamashita, 2014). The difficulty in comparing the results from the L2 studies lies in the variability in the tasks and between the participants. Despite claims that inefficient lexical access would affect reading comprehension more in the beginning stages of literacy (Durgunoğlu, Nagy & Hancin-Bhatt, 1993; Grabe, 2009; Jenkins, Fuchs, van den Broek, Espin & Deno, 2003), only very few empirical studies have been conducted with young L2 learners (cf. Hulstijn, 2011). Dufva and Voeten (1999), however, followed 160 Finnish-speaking schoolchildren longitudinally during Grades 1–3. They found that the strongest single predictor of L2 proficiency at Grade 3 was the level of L1 word recognition speed at Grade 2. In a study by Van Gelderen, Schoonen, Stoel, De Glopper and Hulstijn (2007), the speed of L2 word recognition predicted a small, but significant amount of variance in L2 reading comprehension in Grade 8, but not in later grades (9 and 10).

However, Grabe (2009: 23) states, "Fluent reading comprehension is not possible without rapid and automatic word recognition of a large vocabulary". This suggests that the efficiency of lexical knowledge is more crucial in L2 than in L1 reading comprehension, as cognitive capacity is more taxed in less-automatized language (cf. Verhoeven, 2000). As, for example, Walczyk and colleagues (2004) have advanced for L1, poor readers' processes are mostly as accurate but not as fast as those of efficient readers, which they suggest results from the use of more controlled than automatic processes. A similar pattern has been demonstrated for L2 readers when compared to L1 readers (Shaw & McMillion, 2008). Attention and executive functions have only rarely been studied in relation to L2 processing, but they are proposed to be associated with more developed language skills (Jarvis et al., 2014; Segalowitz & Frenkiel-Fishman, 2005). For example, Bialystok et al. (2008) discuss that a word retrieval task (Letter Fluency) requires executive processes to be performed efficiently, and this is supposedly easier for bilinguals than for monolinguals. Therefore, even though attention-control mechanisms are language-independent, it appears that the extent to which they can be applied when using L2 might be partly dependent on one's proficiency level. This relates to the theories of how and when language-related skills may transferred from L1 to L2, which is discussed further in **Chapter 2.3**.

2.2.3 Fluent lexical access in writing

The influence of fluent lexical access on writing skills is an under-researched area. To measure cognitive processes in writing, the method most often used is a think-aloud protocol, with participants explaining their process either at the time of the writing or afterwards. Other methods include keystroke loggings and pause-time measurements, for instance (see e.g. Manchón et al., 2007; McCutchen, 2000). However, Hayes's (1996) model of **L1 writing** considers that writing consists of planning and formulation in the context of the task environment. This requires a smooth operation of cognitive skills, such as working memory. The importance of fluent lexical access is highlighted at all levels of sub-processes (see also Manchón et al., 2007). McCutchen (2000) reviewed a wide sample of studies dealing with L1 writers and suggested that "without fluency writers cannot move beyond working memory limitations" (p. 16). In a previous study by McCutchen, Covill, Hoyne, and Mildes (1994, cited in McCutchen, 2000), the researchers found that the better writers were the ones who also accessed words in the lexicon more fluently. As for writing in L2, the processing limitations seem to be even more restricting, which contributes to making the writing process much more fragmented compared to L1 (Manchón et al., 2007). Sasaki (2002) found that expert L2 writers wrote longer texts and significantly faster than beginner L2 writers. The novice writers' slowness resulted from pausing to translate L1 words into L2, which interrupted their writing process. Kuiken, Vedder, and Gilabert (2010) report that the correlations between assessments of communicative and linguistic proficiency were higher for the more proficient L2 writers. They propose that the less proficient L2 writers concentrated merely on either the communicative or the linguistic aspects due to memory restrictions, although that is just one of the possible explanations which the researchers hypothesise.

Only a few studies have measured the relationship between fluent lexical access and L2 writing skills. Harrison and colleagues (2015) did find naming speed to explain variance in L2 writing scores, although not in L1 writing. As has been proposed in relation to reading (Van Gelderen et al., 2007), it may be that the differentiating power of fluent lexical access decreases after enough automatization. Schoonen et al. (2002) showed that the speed of lexical access correlated substantially with writing proficiency for children aged 13–14; it did not, however, explain writing proficiency after grammatical and metalinguistic knowledge were considered. The lexical access task used was typing the first letter of stimulus words, and it may be that written fluency does not reflect cognitive fluency as well as an oral task might, and it does not require accurate whole-word retrieval as writing usually does (cf. Fuchs et al., 2002; Georgiou et al., 2013). Furthermore, Sumner, Connelly and Barnett (2013) showed that their dyslexic participants were slow writers not because of motor difficulties, but because of more pausing while looking for words. Speed of word retrieval in writing has, as well, been shown to be possible to train by Snellings et al. (2002). The use of the trained words increased after the training period, and researchers interpreted this as a sign of reduced cognitive load of working memory resulting from easing the process of lexical retrieval. In a later study, the same researchers trained their students on a set of words and found that the use of these words increased in the narrative texts with easier and more fluent retrieval (Snellings et al., 2004). This did not, however, improve the quality of the texts. The role of executive functions in relation to writing has rarely been studied, but Drijbooms, Groen, and Verhoeven (2015) report that composite measures of several tasks tapping into inhibition and monitoring contributed to text length in L1, that is, to the word-level text generation. They argued this was due to success in suppressing irrelevant representations, in selecting the appropriate words, and in actively monitoring one's progress. This is probably more challenging and taxing in L2. To conclude, there seems to be tentative agreement in the literature that fluency of lexical access in some way affects L2 writing, but exactly why and how much is still debated.

2.3 The relationship between L1 and L2 skills

Considering the relationship between L1 and L2 fluency, no easy answers are available. It is evident that knowing one language beforehand will in some respect affect the one(s) learned later, but the nature of this relationship is much discussed. One influential (although oft-debated) theory of the relationship is the concept of **transfer**. Koda (1996) suggests in her review that L1 processing skills seem to transfer to L2 to a certain extent, but the mechanism is still unclear. In addition, the phase of transfer remains unclear, that is, whether it is something that happens in the beginning of language learning when L2 skills are underdeveloped (Cheung & Lin, 2005), or whether there is a **threshold** that is required of either the L1 (Cummins, 1991) or the L2 knowledge before skills can be transferred from L1 (Alderson, 1984; Hulstijn, 2011). Koda (2008: 78), for example, defines transfer as "an automatic activation of well-established first-language competencies, triggered by second-language input", and, crucially, this connection does not stop even with developing L2 skills.

One of the most well-known theories concerning transfer is Cummins's (1991) interdependence hypothesis. Its main claim is that L1 and L2 reading skills are a shared ability and L1 reading skills may be transferred to L2. This requires, however, well-developed automaticity and fluency from L1 (see also Grabe, 2009). For example, the findings of Van Gelderen and colleagues (2007) offer support for this view. On the other hand, Jeon and Yamashita (2014) conducted a meta-analysis of 58 studies studying predictors of L2 reading, and concluded that L2 comprehension correlated more with L2 specific components than with general cognitive measures. Cummins's theory has been criticised for disregarding the effect of L2 proficiency, and a more componential view of the transfer is called for (Verhoeven, 2000; Geva & Ryan, 1993). Therefore, the discussion has moved on to the issue of which domains of language are possible to transfer. In question is whether it is the lower-level skills (such as word recognition fluency) that may transfer or whether it is only the more general cognitive processing ones (such as reading strategies) that may do so (Cheung & Lin, 2005; Hulstijn, 2011; Koda, 2008). Grabe (2009: 144), for example, summarises the current consensus of the transferrable domains as being phonological awareness, decoding, reading strategies, metacognition, and pragmatic knowledge; and the non-transferrable domains as being vocabulary knowledge, morphosyntactics, listening comprehension and orthographic processing.

Nevertheless, the division is perhaps most fruitful if drawn between language-dependent and language-independent processes. Durgunoğlu and colleagues (1993) report that L1 phonological awareness (a low-level cognitive skill) explained variance in L2 word recognition scores. In Koda's (1996) review, she concludes that orthographic and visual strategies seem to be transferred as well. For lexical access measurement, most often only L2 tasks have been used in explaining L2 proficiency, but some results do suggest that speed in L1 lexical access may explain at least some variance in L2 measures (Dufva & Voeten, 1999; Gholamain & Geva, 1999). Bernhardt (2005) proposes that there may be a threshold after which L1 compensates for lack of automaticity in L2 (see also Shaw & McMillion, 2008). Walter (2008) states that in reading, it is not the language-related reading skills per se that transfer, but the ability to apply the more general linguistic and cognitive skills which underlie all language use. When writing in L2, the writers seem to use L1 resources and strategies for problem-solving to a great extent in all stages of composition, especially at the low proficiency levels, and this use is proposed to decrease somewhat alongside the developing L2 skills (Manchón et al., 2007). In reading, even good L1 readers seem to revert to inefficient reading strategies (e.g. pausing, rereading, looking back) when facing time pressure (Walczyk et al., 2004) or when reading in L2 (Alderson, 1984). Regarding L2 writing, beginning writers have been found to translate more from their L1, rather than to spend time on refining their expressions (Sasaki, 2002; cf. Manchón et al., 2007). Harrison et al. (2015) found that the writing scores for English L2 speakers were partially explained by the language-dependent syntactic measures, even though to a lesser extent than for L1 speakers. This was interpreted in favour of the threshold hypothesis. However, caution is called for when considering what and how much can be transferred, as some researchers point out that the similarities and differences between the L1 and L2 affect transfer as well (Geva & Siegel, 2000; Koda, 1996; contrasting results, Jeon & Yamashita, 2014).

Moreover, it is considered that for L2 proficiency, knowing the language is in itself not enough, but also fluency in its use is required (Hulstijn, 2011; Koda, 1996; Segalowitz, 2000). L2 fluency, however, seems to relate to L1 fluency and speaking styles to some extent. For example, De Jong et al. (2015) have stressed that there is a need to determine the L1 baseline first. Their findings showed that the L1 self-corrections explained the substantial amount of variance in the L2 self-corrections ($R^2 = .46$), and the L1 repetitions explained the variance of the L2 repetitions ($R^2 =.43$). This means that L2 fluency seems to be determined partly by L2 proficiency, and partly by L1 fluency, including individual speaking styles and preferences. Kormos (1999), for example, reports that L2 speech rate was significantly influenced by individual speaking style. Therefore, a notable point is that L1 and L2 fluency and lexical access are to some extent related via the general cognitive skills and styles. This connection should be borne in mind to avoid the over-idealised notions of the somehow 'perfect' L1 readers, writers or speakers (Hulstijn, 2011; Lennon, 1990). As Cook (2002) proposes, the influences may be bi-directional in the sense that the new languages influence the L1 use, distinguishing multilingual L1 users from monolingual ones. Therefore, examining and controlling for the L1 fluency behaviour is an important aspect when one studies L2 fluency from the cognitive perspective.

2.4 The aims of the study

The objective of the current study has been to examine L2 fluency from a cognitive perspective; and fluency of lexical access was selected as a possible tool for this. The aspects of cognitive fluency, following Segalowitz's framework, were 1) efficiency (speed and accuracy) of lexical access and 2) attention-control processes. As mentioned in **Chapter 2.1.2**, stability measurements could not be included. The choice of method, lexical access, was proposed to minimise the utterance fluency aspects that may reflect more speaking skills. To control for the individual speaking styles, L1 lexical access fluency was measured in terms of its role in L2 fluency and L2 literacy. The influence of perceived fluency, which relates more to the listener than to the speaker's mental operations, was controlled for by comparing the fluency measurements to the L2 reading comprehension and writing skills of different-level L2 learners. The areas of interest explored in each article are presented in Table 1.

TABLE 1 The distribution of research questions covered in each article.

	Article I	**Article II**	**Article III**
Aspect of cognitive fluency	Efficiency: speed	Efficiency: accuracy	Efficiency & attention
Lexical access	x	x	x
Proficiency	x	x	
L1	x	x	x

3 DATA AND METHODS

In the following section, I present an overview of the participants, the tasks, and the analyses included in the articles of the thesis. The data are a part of the DIALUKI project (Diagnosing reading and writing in a second or foreign language, 2010–2013, the University of Jyväskylä, www.jyu.fi/dialuki). The project was chaired by Professor Charles Alderson (of the University of Lancaster) and Professor Ari Huhta (of the University of Jyväskylä), and was funded by the Academy of Finland, the University of Jyväskylä, and the UK Economic and Social Research Council (ESRC). Several researchers and assistants participated in the project; therefore, I first briefly describe both the overall project and the extent of my own contribution. Next, I describe the lexical access tasks and the measures of literacy skills that were used in the current study. More detailed information about the items and their descriptions are found in Articles I–III. Rating methods of the different tasks is described shortly; it must be pointed out that I was responsible for the lexical access task ratings, but did not participate in the literacy assessments, which were conducted by trained, experienced raters. In the last section, I clarify the distribution of the tasks on which the results in the individual papers are based.

3.1 Participants and data gathering

Participants (N=901) were recruited from 111 schools in Finland, from rural, suburban, and urban areas. The data gathering took place in a cross-sectional setting, and altogether five researchers and 25 assistants were involved in this during the autumn term 2010 and the spring term 2011. I participated in the collection of data by testing 49 participants in both group and individual sessions. The participants were from two language backgrounds: (1) Three age groups of **Finnish-speaking** children learning English as a foreign language participated in the study. They were attending Grade 4, Grade 8, and Grade 11 (second year in upper secondary school, see **Table 2**). All the children had studied English as their first foreign language since Grade 3; therefore, the performances at Grade 4, 8, and 11 were deemed to be reasonable estimates of different levels of L2 proficiency (1, 5, and 8 years of English studies, respectively; more on this rationale in Article II); and (2) Two groups of **Russian-speaking** school children, who were immigrants living in Finland and studying Finnish as a second language, were studied. They were from two school levels: one group from primary school (Grades 3 to 6) and one from lower secondary school (Grades 7 to 9). In the present study, only the results from the primary school level Russian speakers were included, as the older children's results were very difficult to include in the statistical analysis due to their heterogeneity (therefore, in this thesis the final N = 821). **Table 2** presents an overview of the participants who were included in the current thesis (for more information on the participants included in each article, see **Table 4** and **Chapter 4**).

TABLE 2 Summary of the participants of the current thesis.

	N (=821)	**Age**	**Females %**	**L1**	**L2**
Grade 4	210	9–10	52	Finnish	English
Grade 8	208	13–14	56	Finnish	English
Grade 11	219	17–18	53	Finnish	English
Russian primary	184	8–13 (10.9)	52	Russian	Finnish
- preparatory	27				
- regular class	157				

The Russian-speaking groups were further divided into two sub-groups, based on their level of L2 Finnish and length of residence in Finland. The first subgroup of these children was participating in preparatory classes, and the second sub-group was already integrated within the mainstream education in Finnish schools.[1] In Finland, all children of compulsory school age (i.e. 7–17) who are not proficient enough for regular Finnish language studies are entitled to preparatory classes to enhance their Finnish skills and cultural knowledge, irrespective of their grade level. This usually takes about a year, after which they can be integrated into the mainstream classes. The language skills of the Russian speakers in both L1 and L2 were varied; for some, the *de facto* stronger language was Finnish (this issue and its influence on the results are discussed in Article I). Consents to participate in the study were collected from the county, the schools, the parents, and the children.

3.2 Materials

3.2.1 Lexical access tasks

The lexical access tasks included three oral and two written assignments. All except one of the written assignments were performed in both L1 and L2. The oral tasks were part of a larger cognitive test battery (altogether 45–60 minutes), which was conducted individually to the participants by trained assistants. Testing was done with Cognitive Workshop software (the Finnish version was developed by the University of Dundee and the Jyväskylä Longitudinal Study of Dyslexia, see Lyytinen et al., 2004; the English and Russian versions of the tasks were developed by analogy in DIALUKI). The pre-recorded instructions were played to the participant via headphones, and responses were recorded by microphone and saved to the computer for later analyses. The L1 and L2 versions were counterbalanced to minimise the language order effect.

Rapid Alternating Stimulus (RAS) is a word retrieval task measuring naming speed of familiar vocabulary. The version used in the current study consisted of 50 units, arranged in five rows in a semi-random order, and included numbers, colours, and letters (from Ahonen, Tuovinen & Leppäsaari, 2003). In L2, Grade 4 and the primary school group had a slightly easier and shorter version, consisting of 30 units with numbers, colours, and easy everyday objects. The matrix was different because picture naming was considered to be more automatized at the primary school level than foreign language letter naming (see e.g. Denckla & Cutting, 1999; for more on the difficulties in the letter naming, see Article III). The time taken to perform the task was measured by a stopwatch, and was converted to time-per-unit used as the speed measure. The accuracy of the items was analysed from the sound files recorded during the tasks, as a percentage of accurate and fluent items (Article II), and according to the categorisation based on the processing limitations in fluency (Article III; explained in more detail in **Chapter 3.3.2**).

In the current study, a RAS task was used since, in addition to the speed of word retrieval required in the basic naming speed task (RAN, Rapid Automatic Naming: Denckla & Rudel, 1976), there is also the element of shifting between different semantic categories. Thus, it is more challenging and may tap more directly into reading comprehension (e.g. Närhi et al. 2005). The RAS task was originally designed by Wolf (1986); she found it significantly differentiated average and dyslexic readers from each other already in kindergarten. RAS tasks which included letters, numbers, and colours especially contributed to reading comprehension at all levels (see also **Chapter 2.4.1**). Wolf interpreted this as resulting from the extra cognitive load of a RAS task as compared to a RAN one, in that a RAS task requires the integration of the lower- and higher-level reading processes. Närhi et al. (2005) found support for this view, stating that adding the set-shifting accounted for variance that was not covered by naming the individual RAN items. This set-shifting would, therefore, measure a process more closely related to reading, with attention given to inhibition in addition to the simpler letter or object naming.

Rapidly Presented Words (RPW) is a task measuring perceiving and reading words as wholes. In the current version, first there was a prompt on the computer screen, then the target word was shown for 80 milliseconds, after which it was replaced by a mask (non-letters such as #%&:) to erase the visual image from the screen. The child had to say the word aloud or confirm not having perceived it. All the words were very frequent, but increased in length from two letters to nine letters (with 8–14 words in total). Most words were simple forms, but some of the Finnish words were inflected ('koulu*n*' – 'school+*GEN*'). The score from the task was a percentage of the correct answers; the speed and accuracy components could not be distinguished. The RPW task is an often-used measure for sight-word reading, as the exposure times are usually too short to allow a complete access to one's lexicon (see e.g. Booth et al., 1999). Sight-word reading is considered an essential phase of mature reading skills (Ehri, 1991). When this is automatized, as Cattell (1886) already found, fluent readers are quicker in naming whole words than they are in naming individual letters (later referred to as word superiority effect, see e.g. Balota et al., 2006).

The Word List task consisted of 105 words arranged in three columns, and the children were instructed to read aloud as many words and as accurately as they could in 60 seconds' time. The task required two types of reading processes. The beginning of the list included short and more frequent words that could be read utilising a method tilting toward sight-word reading; whereas the gradually longer, more complex words towards the end required more accurate decoding skills. The speed score of the task was counted as syllables per second, and the accuracy of the items was analysed in a similar manner as in the RAS task. The Finnish word list used in the current study was a part of the standardised Lukilasse test battery (Häyrinen, Serenius-Sirve & Korkman, 1999); for the English L2 and Russian L1 versions, the word lists were designed by the DIALUKI project by sampling from frequency lists.[2] Consequently, it was not possible to produce exactly equivalent word lists in each language, but matching the difficulty level of the versions was attempted. Nevertheless, even though the number of words was equal, the lists differed in the number of syllables (Finnish 379 syllables; English 183 syllables; Russian 299 syllables). Finnish words included affixation and complex inflections, and the longest Finnish words were 22 letters long (e.g. prosessikirjoittaminen 'process writing'). As Grabe (2009: 120) points out concerning reading rate, English is, as a language, faster to read than the more information-packed languages like Finnish. Due to this, the Finnish word list proved to be quite challenging; on the other hand, Aro & Wimmer (2003) have suggested that the reliable phoneme-grapheme-connections and transparent orthography make decoding the words easier.

When reading a list of isolated words, the reader does not receive any aid from the syntactic or semantic context. Particularly Fuchs and colleagues (cf. 2001) have argued that passage readings are more indicative of further reading skills, and Jeon and Yamashita's (2014) meta-analysis found the correlation between these measures to be very high (.92). This result, however points to a major overlap between the measures, as the authors conclude as well, and how informative this is in the end is an interesting question. Geva, Wade-Woolley, and Shany (1997) reported that isolated L1 word reading distinguished between good and poor readers in both L1 and L2; a further interesting finding was that in L1, reading a text was faster than reading isolated words, but in L2 this was not the case. Furthermore, Jenkins et al. (2003) reported that the more advanced readers benefited more from contextual cues than did the less advanced ones. Therefore, it seems that poor L1 readers and beginning L2 readers are similar in that neither benefit from the contextual cues, and reading an isolated word list is in some way a more equal task in revealing more of the cognitive difficulties related to reading. On the other hand, when the reader is under a time constraint, good decoding skills may help in the comprehension process (Walczyk et al., 2004). Thus, the reading-the-list format may be able to assist researchers in tapping into more general cognitive fluency.

The **Word Chain** was a paper-and-pencil task with a short text written without punctuation or spaces between the words (adapted from Holopainen, Kairaluoma, Nevala, Ahonen & Aro, 2004). The task measures silent reading skills, and speed and accuracy of word recognition (see e.g. Georgiou et al., 2013). It requires decoding, but sight-word recognition aids the speed in the task. The task was to separate the words by pencil strokes at word boundaries. An example of one line in a Grade 4 English (L2) text is:

sothenextdaythethreelittlepigslefthomethefirstpigmadeahomefromstrawthesecondpig

The excerpts were different in L1 and L2, and in different grades so as to achieve more age-appropriate tasks. The texts originated in the language course books, and were used with the kind permission of the publishers.[3] The speed score was the time to conduct the task to the nearest 30 seconds, and the accuracy score was the percentage of the correct answers from the total number of words. The testing situation included time pressure, and this was hypothesised to induce some trade-off between the speed and accuracy. Test-retest reliability for the original Finnish version of Word Chain task is .70 and .84 in two samples (Holopainen et al., 2004). Miller-Guron and Lundberg (2000) reported that Word Chain task differentiated significantly dyslexic students from normal readers in both L1 and L2. Akamatsu (2008) used a Word Chain task with de-contextualised words in his training study, and found that even this kind of out-of-context training improved participants' word recognition fluency. In the current version, short story excerpts were used so that in addition to involving efficient access to one’s lexicon, the task required some semantic considerations; thus, it was not a pure word recognition task.

The **spelling error detection** task was a written task, and it was conducted in the classroom for the whole class (Holopainen et al., 2004). It was carried out only in L1 Finnish, and only for the Grade 8 and Grade 11 students, as the task was designed for young adults. The task comprised 100 Finnish words on paper, each containing a spelling error. The spelling errors were of different categories: a missing letter, an extra letter (*lauvantai – lauantai* 'Saturday'), a wrong letter (*heunäkuu – heinäkuu* 'July'), or an incorrect morpho-phonemic variation (*aurinkonkukka – auringonkukka* 'sunflower'). Children were instructed to mark all the spelling errors they detected but were not to correct them. The number of correct markings out of 100 in a period of 3.5 minutes was calculated; therefore, the speed and accuracy component could not be distinguished.

3.2.2 L2 literacy skill measures

L2 literacy skills were measured by reading comprehension and writing tasks. These served as the dependent measures to compare if fluent lexical access is reflected in the language proficiency of the children assessed in the different age and language groups. The tasks were conducted in a group situation for the whole class at the same time. All writing tasks and most of the reading tasks were paper-and-pencil style. For Grades 8 and 11, some digital-format reading tasks from Dialang (an online diagnostic system, see Alderson & Huhta, 2005) were used. The specific items are explained in Articles I–II.

Reading comprehension: The texts chosen for reading comprehension assessment were age-appropriate, and they were designed to measure different dimensions of reading: the comprehension of both main points and details, as well as inference-making. The instructions for the tasks were provided in both L1 and L2 to ensure that everyone understood clearly what they were expected to do. **Grade 4** tasks were, first, matching sentences from a story to appropriate pictures; and second, choosing words from a bank of single words to complete a short cloze text (Pearson Young Learners test, level A1 to A2).[4] The task for **Grade 8** included four texts from the Pearson Test of English General (levels A2 to B1), and 30 items from Dialang (intermediate level). The answer types included multiple-choice tasks, short open answers, and gap fills. In **Grade 11**, there were five texts from Pearson Test of English General (levels B1 to B2), and a further 30 items from the Dialang upper-intermediate level with similar answer types. For the **L2 Finnish** learners in preparatory classes, three short texts from the National Certificate for Language Proficiency were used[5]; the more advanced L2 Finnish speakers answered questions from ten short texts in Dialang that were considered suitable for young children, from elementary to intermediate levels (roughly A1 to B1).

Writing tasks: The tasks originated from Pearson testing and from the Cefling project (University of Jyväskylä, 2007–09, see Alanen, Huhta, Jarvis, Martin & Tarnanen, 2012). The topics addressed aspects of personal life, either at home or school. The **Grade 4** writing task was "Answer your friend's message and tell of the things you like". Some prompts were given to the participants, including a reminder to greet the friend appropriately and to mention at least one colour, food, and song or artist they like, but the amount of text was not determined. The participants from **Grade 8** wrote two compositions. The first one was entitled "How do you travel to school?" with instructions to argue for your choice, and discuss the pros and cons of the chosen method and possible better choices (100–150 words). For the second composition, the students could choose from two topics, "Mobile phones in schools" or "Boys and girls should be in different classes", and write an opinion piece. The instructions were to write at least a couple of sentences for or against the argument. **Grade 11** students' first writing task was an article entitled "A journey I'll never forget" to be written for a fictitious travel magazine, with instructions to include where the trip happened and why it was special (100–150 words). The second was the choice of the same opinion pieces as in Grade 8 with same instructions. Beginning **L2 Finnish** speakers wrote on the topic "Tell your friend of the things you like", with the same design as that of Grade 4. The more proficient L2 Finnish students wrote on the topic "Tell us about something funny or scary that happened to you", with further questions pertaining to describing what happened and explaining their feelings with at least a couple of sentences.

3.2.3 Rating of the tasks

In the case of the lexical access tasks, the time taken to complete them and the percentages of accuracy were coded by ten DIALUKI assistants and double-checked by the author. Certain concessions were made in the RAS task; for instance, an answer was considered correct if a child named *house* as *home*, *pencil* as *pen* or *blue* as *violet*, as long as the label was consistently used throughout the task. Colour blindness was controlled for in the same manner. For the Finnish and English data, an additional coding of inaccuracies in the lexical access tasks was designed (this is explained in more detail in **Chapter 3.3.2**). As the accuracy analyses could not be dichotomously correct or incorrect, 10% of the data from each of the Finnish L1 groups were double-checked by a second rater instructed on the assessment criteria by the author. The overall agreement rate on the location and type of inaccuracies was 95.6%.

The **reading comprehension** tasks were rated by two trained researchers, and as most of the questions were of multiple choice format, they were fairly objective to score. The reading test scores were analysed with Winsteps, a Rasch analysis programme, which produces an interval scale logit measure of the reading skill for each learner that was used as the dependent variable in the analyses. The rating of the **writing** tasks was based on Cefling methodology (see Alanen et al., 2012). The Common European Framework of Reference (CEFR, Council of Europe, 2007) and the Finnish National Core Curriculum scales (the official reference scale in the Finnish school system[6]) were jointly used. The communicative CEFR scales were employed to aim at assessing a child's ability to use the language rather than to exhibit certain linguistic features. A total of 11 raters for the rating of English and nine for Finnish from the University of Jyväskylä were employed by DIALUKI. The raters participated in a training session, and each one assessed 100–500 scripts; furthermore, each text was rated independently by two raters. The overlapping rating design allowed the data to be analysed by Facets (a multifaceted Rasch analysis program). The Rasch analysis produces proficiency values (similar to those described above for reading scores) that take into account both the raters' leniency or severity and the task difficulty from different perspectives. The advantage to this kind of rating system is argued to be the possibility to use the ratings across several tasks and raters, which should result in more objective assessments when the task type and rater subjectivity effects are reduced (Alanen et al., 2012). The facets used in these analyses included the ratings for L2 learners, the raters, and the tasks. The results from these ratings were used as the dependent measure of L2 writing.

3.3 Analyses

3.3.1 Statistical analyses

The statistical analyses in the articles were performed by IBM SPSS Statistics 22 software. Means and standard deviations were counted for the speed and accuracy in lexical access for each age and language group. In **Article I**, the performance of the different age and language groups was compared with the analysis of variance (ANOVA) to see if the groups could be distinguished by the speed of the lexical access in L1 and L2. A set of stepwise multiple regression analyses were conducted separately for each group (Grades 4, 8, 11, and Primary school) to see whether the speed of lexical access, in both word recognition and word retrieval, and in L1 and L2, could explain the variance in L2 reading comprehension and writing results.

In **Article II**, correlations and hierarchical regression analyses were calculated separately for all three Finnish-speaking grades. The main interest was to examine whether in addition to the speed of lexical access, the accuracy of lexical access in both L1 and L2 attributed to L2 literacy scores. To tease out the unique contribution of accuracy (as mentioned above, often overlooked) the speed variables were controlled for first, by entering them in the model before the accuracy variables. In the same manner, L1 results were controlled for before the L2 measures to evaluate the effect of L2 efficiency after the supposedly more general efficiency measured by L1 was controlled for. Scores from the naming task and word recognition tasks were all entered into the model in separate steps to assess the unique contributions of word retrieval and word recognition. Accuracy scores for RAS L1 and L2 tasks, and for the Word Chain L1 task were too skewed, and could not be included in the analyses; thus, only the accuracy percentages for Word List L1 and L2, and Word Chain L2 were used. There were altogether six steps, performed in the following order: naming speed L1 (RAS), naming speed L2, word recognition L1 (Word List and Word Chain), word recognition L2, word recognition accuracy L1 (Word List), and word recognition accuracy L2 (Word List and Word Chain).

3.3.2 Accuracy analysis

A categorisation of inaccuracies based upon the cognitive fluency construct and processing limitations in language use was designed for Article III. Inaccuracies were possible to be analysed separately in the RAS and Word List tasks. First, the instances of each type of inaccuracy were counted, and their proportions were compared across the grades. As the number of words read in the Word List task differed according to each individual and language, and in the RAS task according to grade and language, to be able to compare the inaccuracies among grades and languages, the relative proportions of inaccuracies were calculated (the amount of inaccuracies in each category by the total number of words read / named for each age-group separately, cf. Protopapas et al., 2013). Third, χ^2 were performed on the counts to see if the proficiency levels differed significantly in any categories.

Based on Segalowitz's (2000, 2010) concept of cognitive fluency, I proposed that the reasons for inaccuracies might stem from two sources, both materialising as different inaccuracy phenomena (see Fluency model in **Chapter 2.3.3**). First, the inaccuracies were proposed to originate from **inefficiency**, manifesting itself as lack of speed or accuracy, being visible in lower-proficiency L2 levels, and decreasing when proficiency increases. (See **Table 3**. N.B. The numbers enclosed in parentheses in the text that follows correspond to the itemised Table 3 content.) Fluent reading requires both fast and accurate word recognition, and, particularly in English, most words must be recognised as wholes (Ehri, 1991). When this does not happen efficiently, this may lead to (1) guessing words based on their form. The result is often a more frequent word (see e.g. Balota et al., 2006), and this phenomenon is proposed to be more pronounced in word recognition than in word retrival (Gollan et al., 2011). The connectionist models (Seidenberg & McClelland, 1989) also include the possibility that the frequently used (more automatized) words inhibit the activation of less-used words. Next, inefficient lexical access may also lead to (2) skipping items, either by accident or by actively avoiding difficult words. Furthermore, in the RAS task, inaccuracies may materialise in (3) the within-category errors, where the item is replaced by another item of the same category. In addition, a category of (4) pronunciation problems was included with otherwise uncategorisable items, miscellaneous stammerings, etc.

TABLE 3 Inaccuracy categories in lexical access, based on the cognitive fluency construct, and their descriptions.

Category	Description
Efficiency	
1. Guessing	producing wrong responses that are graphemically similar to the target words, frequency-based (Broadbent, 1967; Balota et al., 2006)
2. Skipping	skipping items either accidentally or strategically (Marian et al., 2013)
3. Within-category (RAS)	producing wrong responses that are within the same category, such as naming 6 as 7, A as E, or green as blue
4. Pronunciation problems	stammering when encountering difficult phonemes or combinations
Attention	
5. Self-correction	correcting one's own speech requires attention and cognitive resources (Segalowitz, 2010)
6. Repetition	repeating items, subject to individual differences and strategic use (e.g., Kahng, 2014; Peltonen & Lintunen, 2016)
7. Inhibition difficulties	difficulty with suppressing irrelevant information are proposed to be connected with more advanced skills as the automatic activation is stronger (Marian et al., 2013; Jones et al., 2008)
8. Set-shifting (RAS)	set-shifting is more difficult with stronger activation and may be easier with less proficiency; e.g., switching from activated letters to numbers (Segalowitz & Frenkiel-Fishman, 2005)

The second category of inaccuracies included the ones associated with **attentional processes.** First were the (5) self-corrections, which require allocation of attention to the accuracy of the output. These were, therefore, hypothesised to be found more for the more proficient languages/users, as monitoring one's own speech requires resources that low-proficiency language users often do not possess (see e.g. Kormos, 2000; Walczyk et al., 2004; also, Schmidt, 1992). (6)

Repetitions were included in this category as well, following Skehan's (2003) framework, as they are considered to belong to repairs of speech and are possibly sometimes used intentionally as a strategic device requiring monitoring of speech (e.g. Peltonen & Lintunen, 2016). Further, the inaccuracies resulting from (7) inhibition difficulties were included. These were proposed to result from the inability to inhibit irrelevant information in the form of pre-activation of items in the visual field close to the target word (Jones et al., 2008; Simola et al., 2009). This was hypothesised to be found for the more automatized languages. The last category was (8) set-shifting difficulties, as in the RAS task, sometimes it was expected to be difficult to switch from an activated category to a new one. An example of this might be when a number is followed by a letter; for example, the number 'six' is followed by the letter 'S', and the participant names the letter erroneously as the number 'seven'; that is, s/he is unable to switch from the activated category of numbers. This proposition is based on the notion that automaticity may be revealed in the stronger language by stronger activation of vocabulary (Marian et al., 2013). It should be noted that the RAS and Word List tasks, by means of which the inaccuracies were studied, are not designed to reveal these last two types of mechanisms. Consequently, the results one might expect were at best tentative and should later be corroborated with other tasks designed to measure the executive functions (e.g. Flanker, Stroop, or Simon Tasks, see e.g. Diamond, 2013).

3.4 Organisation of the articles

Articles I–III are all based on the same data set, but the participants and materials included in each study differed slightly, depending on the research questions and restrictions posed by the data. The distributions of the materials in each paper are presented in **Table 4**. The Russian speakers were included only in Article I; their results were too heterogeneous to be used in the more rigorous statistical analyses of Article II. In Article III, only the Finnish-speaking students were included, as the interest lay in comparing clearly distinguishable proficiency levels and related accuracy performance. The Russian-speakers' performance in L2 tasks fell somewhere in between Finnish Grade 4 and Grade 8, and would have required more individual scrutiny. In Article III, only the RAS and Word List tasks were analysed, because in those, it was possible to categorise the oral inaccuracies. The two other lexical access tasks were written ones, and the RPW task did not provide enough material for analysis.

TABLE 4 Overview of the materials.

	Article I	Article II	Article III
Participants			
Finnish L1	x	x	x
Russian L1	x	–	–
Lexical access			
RAS (L1, L2)	x (S)	x (S)	x (A)
Word List (L1, L2)	x (S)	x (S + A)	x (A)
RPW (L1, L2)	x	–	–
Word Chain (L1, L2)	x (S)	x (S + A)	–
Spelling errors (L1)	x	–	–
Literacy			
L2 reading	x	x	–
L2 writing	x	x	–

Note: RAS = Rapid Alternating Stimulus. RPW = Rapidly Presented Words. S = Speed. A = Accuracy.

4 RESULTS

4.1 Article I: Speed of lexical access and connections to L2 literacy skills

As **speed of lexical access** has been shown to be a good predictor of L1 reading development, the first article concentrated on examining its relationship to second and foreign language literacy skills. Lexical access was measured in both L1 (Finnish or Russian) and L2 (English or Finnish). The age and language groups were compared by using the speed of performance on word recognition and retrieval. The main interest in the paper was to explore whether the speed of lexical access was connected to **L2 reading comprehension and writing performance**. Possible differences among the age groups, as well as between the second language learners (Finnish as a second language, FSL) and foreign language learners (English as a foreign language, EFL), were examined. The design made it also possible to examine whether the performance speed in L1 had any effect on the L2 skills.

In the study, it was found that, first, the speed of lexical access in both L1 and L2 increased with age, as expected. Furthermore, both primary school groups (EFL Grade 4 and FSL primary school) performed faster in the L1 versions of the tasks than in the L2 versions. This is a reflection of the state of their L2 skills in the beginning of development, as the difference between the L1 and L2 in the speed of lexical access decreased with the older students. Speed of lexical access in L1 and L2 was also found to differentiate the groups (ANOVA), except for the Rapidly Presented Words (RPW) task, where the older groups' performance reached a ceiling (for similar results, see Yamashita, 2013). Second, a set of stepwise multiple regression analyses were used to examine to what extent the speed of lexical access in L1 and L2 could explain the variance in L2 literacy skills. The amount of variance explained was quite substantial (see **Table 5**); interestingly, the speed of lexical access (in L1 and L2) explained the variance in L2 writing performance better than in reading comprehension. In addition, the performance of second language learners (FSL) was somewhat better explained by these tasks than the performance of the foreign language learners (EFL). Overall, the FSL learners' performance on the L1 tasks (Russian) was similar to that of the Grade 4 EFL learners (Finnish L1), their closest peers in age. However, concerning speed, the FSL learners' performance in L2 (Finnish) was closer to that of Grade 8 EFL learners in their L2 (English). This highlights the different experience of learning, as the second language learners are more exposed to the second language in their everyday living environment. For some of the children, there is more exposure to the language, but on the other hand, the variability of the L1 Russian skills might have reflected attrition as well.

TABLE 5 The amount of variance in L2 reading and writing explained by speed of lexical access and the first two variables accounting for the largest proportions of the variance. (Pearson correlations to literacy skills.)

L2	Age groups	Dependent variable	Adj. R^2	1st IV	2nd IV
EFL	Grade 4	Writing L2	**27.2 %**	RPW L1 (.491)	RAS L2 (-.401)
		Reading L2	**18.8 %**	RPW L2 (.392)	RAS L2 (-.352)
	Grade 8	Writing L2	**42.2 %**	RAS L2 (-.566)	WC L1 (-.451)
		Reading L2	**21.8 %**	RAS L2 (-.230)	RPW L2 (.298)
	Grade 11	Writing L2	**32.2 %**	RAS L2 (-.459)	Spell L1 (-.249)
		Reading L2	**24.2 %**	RAS L2 (-.399)	WC L1 (-.321)
FSL	PriSchool	Writing L2	**45.9 %**	List L2 (.638)	List L1 (.080)
		Reading L2	**30.0 %**	List L2 (.500)	RPW L2 (.476)

Note. EFL = English as foreign language. FSL = Finnish as second language. PriSchool = Primary school Russian speakers. RPW = Rapidly Presented Words. RAS = Rapid Alternating Stimulus. WC = Word Chain. Spell = Spelling errors. List = Word List.

The age and proficiency level comparisons revealed interesting differences in the distribution of the predictors. For the younger students, the best predictors were the ones that measured efficient **word recognition**, the Rapidly Presented Words (RPW) and Word List tasks. For the EFL Grade 4 group, the largest amount of the variance was explained by the RPW results, in either L1 or L2. Similarly, for the FSL learners in primary school, the Word List results in L2 explained the largest proportion of the variance in both L2 literacy measures.

Contrastingly, for the older EFL groups (Grades 8 and 11), the **word retrieval** (RAS) task in L2 explained the largest proportion of the variance. These tasks measure different aspects of lexical access: word recognition skills build mainly on efficient decoding, whereas word retrieval is more of a direct measure of automatic connections to one's lexicon (see **Chapter 2.2**). This finding suggests that there are possible differences in the reading and writing skill constructs as a person's age and proficiency increase. A further notable point concerning the question of L1 influence is that the effect of L1 was not limited to beginning L2 learners. L1 word recognition tasks accounted for a small independent variance in the regression analysis in both Grade 8 (Word Chain 7.3% in writing) and Grade 11 (Word Chain: 5.0%, and Word List: 4.6% in reading; Spelling errors: 6.8%, and Word List: 5.5% in writing). Word Chain and Spelling errors were somewhat more challenging tasks and their contributions may in part be hypothesised as deriving from the additional syntactic and contextual aspects they measure. It should be noted that in the regression analyses, some related variables may be obscured by others due to overlap, as they partly measure some of the same skills, probably tapping into the more general cognitive mechanisms, in addition to the different aspects of fluency. On the other hand, these relationships between general skills and language-specific skills, and between decoding and lexical access, are also susceptible to age and proficiency development. As a result, similarly to the literacy tasks, the word recognition tasks may not measure the same aspects at different ages. It should be taken into account, furthermore, that the ceiling effect, especially in the RPW task results, may have prevented this task from explaining variance for the older students, and a highly interesting question remains whether its effect could be teased out by a more difficult task setting.

4.2 Article II: Accuracy of lexical access and connections to L2 literacy skills

Definitions of fluency often include both a speed and an accuracy component (Grabe, 2009; Segalowitz, 2010); nevertheless, measurement of accuracy of lexical access has been mostly overlooked, especially in L1 studies (e.g. Bowey et al., 2004; Di Filippo et al., 2005; Kirby et al., 2003). In Article II, it was proposed that for the L2, the individual differences in accuracy might be greater and serve to differentiate learners, as the vocabulary is still in the process of automatization and accurate decoding of words being learnt (following e.g. Ehri, 1991; see also Geva et al., 1997). Therefore, the main question of the paper was whether the **accuracy of lexical access** had a unique explanatory power for L2 literacy scores, independently from the speed of lexical access. For the purposes of this article, only the EFL students were included (N=637). The unique contributions of word retrieval and word recognition were analysed, as these are rarely used in the same context. The accuracy results of RAS L1 and L2, and of

Word Chain L1 could not be included because of the aforementioned ceiling effect. The performances of the three EFL grades (Grade 4, 8, and 11) were examined to see whether there were differences in the effect of accurate lexical access related to the L2 proficiency level. As L1 fluency has been found to explain L2 fluency to some extent (De Jong et al., 2015; Dufva & Voeten, 1999), we expected to see some effect of L1 in this study as well (for the methodology, see **Chapter 3.3.1**).

The accuracy scores mostly increased, and the within-group variation decreased, as the students became more proficient (overall accuracy percentages in both languages for the RAS task were 92–99%; for the Word List task 88–97%, and for the Word Chain task 79–99%). Across all grades, the strongest correlation was found for L2 Word Chain accuracy, in both L2 reading comprehension scores (Pearson r = .56, .55, and .57) and writing scores (.53, .56, and .54). In the regression analyses, the largest individual proportion of the variance in all grades was accounted for by word recognition accuracy in L2. Word retrieval in L1 and word recognition speed in L2 mostly failed to explain any additional variance; however, all the other measures (word recognition speed L1, word recognition accuracy in L1 and L2) offered some small contributions. For example, word recognition accuracy in L1 made a small, but significant contribution in both Grade 4 (6%) and Grade 11 (3%) to the explained variance. As was found in Article I, L2 writing was better explained by fluency of lexical access than was L2 reading comprehension, although the word recognition accuracy added more to the explained variance in reading comprehension than in writing (see Figures 2 and 3).

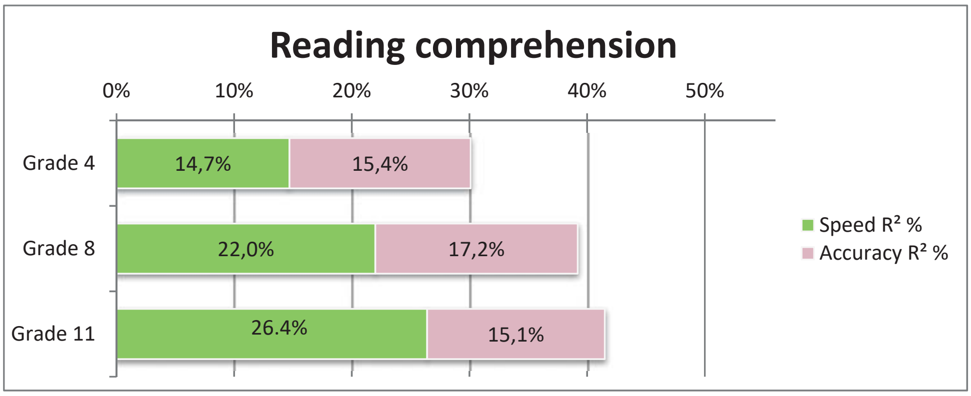

FIGURE 2 The amount of variance in L2 reading comprehension explained by the speed of lexical access in L1 and L2; and the amount of additional variance explained by the accuracy of word recognition (both L1 and L2).

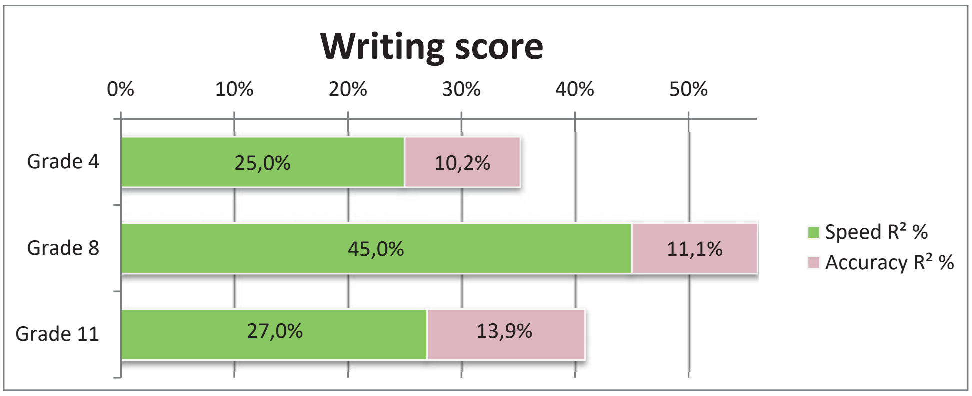

FIGURE 3 The amount of variance in L2 writing skills explained by the speed of lexical access in L1 and L2; and the amount of additional variance explained by the accuracy of word recognition (both L1 and L2).

The contribution of accuracy in lexical access processes in addition to speed was therefore demonstrated. Overall, the L1 and L2 lexical access tasks explained the largest proportion of the variance of L2 reading in Grade 11, and of L2 writing in Grade 8. Therefore, the previous results indicating that fluency of lexical access would be the most decisive factor in the beginning of language learning, both in L1 and L2, were not corroborated for this data set (e.g. Durgunoğlu et al., 1993; Fuchs et al., 2001). For the L1 measures, no clear pattern was found; nevertheless, associations between fluency of L1 lexical access and L2 literacy outcomes were found in all grades. This has a bearing on the discussion of the L1 transfer and thresholds, as, for the most part, these connections did not decrease. This indicates that the influence of L1 skills, or the more general cognitive skills, can be found in very different levels of L2 proficiency. The correlations in the speed in the tasks between the L1 and L2 versions were found to be quite high and/or increase by proficiency; in contrast, the correlations between the accuracy scores were much lower. These results were interpreted in the light of language-dependent versus –independent processes.

4.3 Article III: Inaccuracies in lexical access as reflections of inefficiency and attention-control processes

The third research paper examined the accuracy of lexical access in detail. A fine-grained analysis of accuracy was possible to perform for the results from the **RAS** and **Word List** tasks, and this was conducted for both L1 (Finnish) and L2 (English). The inaccuracies belonging to two cognitive fluency categories were examined in the performance of L2 learners from the three Finnish age groups. The main aim was to try to find connections between the surface fluency phenomena and to the construction of cognitive fluency by examining how the different types of inaccuracies were associated with the various proficiency levels. The hypothesis was that the inaccuracies related to inefficient processing would show a different distributional pattern than the inaccuracies related to allocation of attention. Inefficient lexical access was proposed to be connected to low proficiency levels, and to result in, for example, guessing items. Attention control was proposed to be related to more developed language skills, with, for example, more resources available for monitoring one's own speech. The detailed framework for the inaccuracies is presented in **Chapter 3.3.2**. The categories of inaccuracies in inefficient processing were (1) guessing, (2) skipping, (3) within-category, and (4) pronunciation problems. For attentional processes, the inaccuracy categories were (5) self-corrections, (6) repetitions, and difficulties in (7) inhibition (nearby words affect the parafoveal preview) and (8) set-shifting (e.g. switching from numbers to letters).

First, the total number of inaccuracies was counted. There were altogether 817 occasions of inaccuracies in the RAS L1 results, and 1623 in L2 ones; for the Word List task, the sums were 2758 in L1, and 2092 in L2. These counts reveal first important finding: there is no marked difference in the raw number of inaccuracies between the L1 and L2. Next, the proportional amounts of the inaccuracies were counted. The overall pattern was that in the task with the least cognitive requirements (RAS L1), the attention-control inaccuracies were quite stable across the grades. In the more challenging tasks (RAS L2 and Word Lists), however, their amount increased alongside proficiency in most instances (see Figures 4 and 5). For the inaccuracy types in each task, for the RAS L1 task results, the largest portion of the inaccuracies in all grades fell into the attention-control categories (self-corrections in Grade 4: 35%, Grade 8: 43%, and Grade 11: 43% of all the inaccuracies). The second most common type of inaccuracies resulted from inefficiency in Grades 8 and 11 (within-category errors: 15% and 20%, respectively), but from attention-control in Grade 4 (repetitions: 17%). This most likely reflects the use of repetitions as a strategic device which the older students have no need for anymore in a task as easy as RAS. In the RAS L2 task results, the inefficiency categories constituted the largest proportion of inaccuracies (pronunciation problems in Grade 4: 33%, Grade 8: 73%, and Grade 11: 67%). This was the case in Grade 4 with the second most common type as well (skipping: 29%); however, the ones relating to attentional processes, that is, self-corrections, were the second most common inaccuracies for Grades 8 (14%) and 11 (18%). In the Word List task results, the largest proportion of inaccuracies were related to inefficiency (pronunciation problems) in both L1 (Grade 4: 75%, Grade 8: 62%, and Grade 11: 57%) and L2 (68%, 69%, and 41%, respectively). For Grade 4, inefficiency was the reason behind the second largest proportion (guessing: L1: 13% and L2: 18%). For the higher grades, the second most common inaccuracies in L1 were those relating to attentional processes (self-corrections in Grade 8: 18% and Grade 11: 23%); but in L2, those of inefficiency (Grade 8: 14% and Grade 11: 28%).

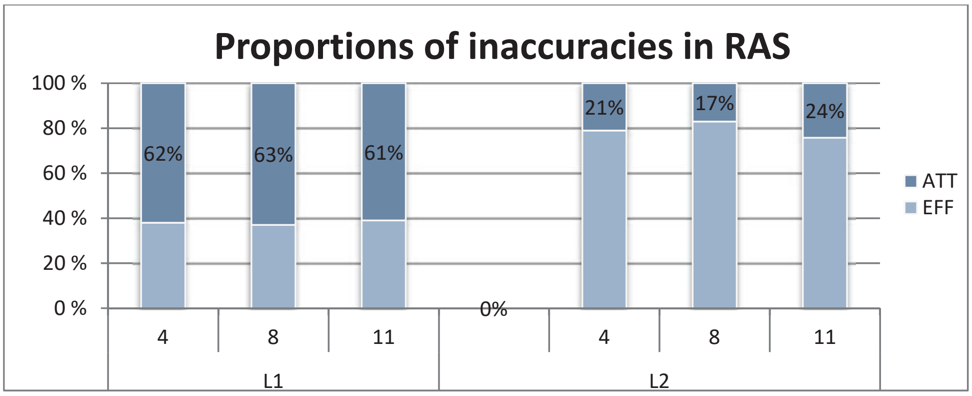

FIGURE 4 Proportions of attention (ATT) and inefficiency (EFF) inaccuracies in the RAS task in each grade in L1 and L2. Attention includes the following: self-corrections, repetitions, inhibition difficulties and set-shifting. Inefficiency includes the following: guessing, skipping, within-category errors, and pronunciation problems.

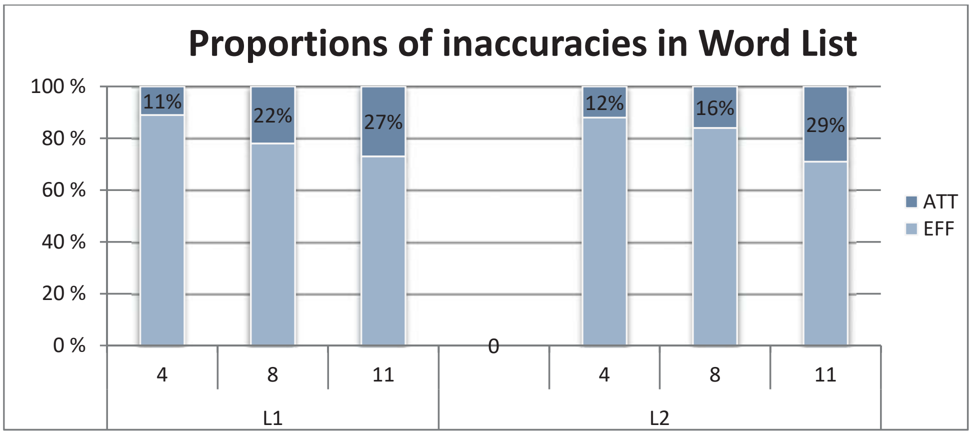

FIGURE 5 Proportions of attention (ATT) and inefficiency (EFF) inaccuracies in the Word List task in each grade in L1 and L2. Attention includes the following: self-corrections, repetitions, inhibition difficulties and set-shifting. Inefficiency includes the following: guessing, skipping, within-category errors, and pronunciation problems.

To compare behaviour among the proficiency levels, χ^2 scores were counted. Significant age group differences were found for the RAS L1 task, which indicated that students in Grade 4 corrected themselves and made repetitions more often than did the other groups. In the RAS L2 task, Grade 4 was left out of the analysis due to task differences, but Grade 8 had more self-corrections, within-category errors, and pronunciation problems. There was also a tendency in Grades 8 and 11 to have more difficulties with set-shifting in L1 than in L2 (Grade 8: L1 17% vs. L2 13%; and Grade 11: L1 19% vs. L2 16%). In Grade 11, furthermore, the students seemed to self-correct more in L1 versus in L2 (45% vs. 36% of the students self-corrected). In the Word List, the χ^2 test showed significant effects for Grade 4, with more pronunciation problems than expected;

on the other hand, concerning self-corrections, there were fewer errors than expected for Grade 4 in both L1 and L2. Both the χ^2 tests and relative proportions hinted that all the groups tended to self-correct more often in L1 than in L2. The self-correction in the L1 Word List task was used in Grade 4 by 39% of the children, in Grade 8 by 57% of them, and in Grade 11 by 61% of them. Comparatively, self-correction in the L2 Word List task was used in Grade 4 by 31% of the students, Grade 8 by 49% of them, and Grade 11 by 49% of them. A similar pattern was found for inhibition inaccuracies (Grade 4: 19% L1 vs. 12% L2; Grade 8: 19% L1 vs. 15% L2; and Grade 11: 23% L1 vs. 21% L2), although these were too few to yield significant effects. These results are, however, in support of the original hypothesis that with lower language skills, either at the beginning of language learning or in a language less developed within an individual, the attentional processes are less available and more difficult to use.

Some additional points are worth noting. There were differences among the tasks in terms of skipping items: the RAS task results in Grades 4 and 8 indicated students skipped more items in L1, but in the other instances, skipping was found to be more prevalent in L2. This was interpreted as that in the RAS task, skipping was mainly due to (too) fast performance; but in the Word List task, skipping (the difficult-looking items) was more due to strategic reasons (cf. Marian et al., 2013, for strategic skipping in the Stroop task). In L1, most of the skipping instances were noticed and corrected; in L2 this was not the case. This offered confirmation to the overall hypothesis that there are more available resources in L1 for speech monitoring and correcting, and only gradually can these be performed in L2 as well (e.g. Kormos, 2000). For repetitions, the groups' behaviour as well as L1 and L2 behaviour were indistinguishable, and the conclusion was that repetitions seem to be more related to personal speaking styles or strategies (in accordance with e.g. Bosker et al., 2013; Engelhardt et al., 2010; Peltonen & Lintunen, 2016).

4.4 Summary of the results

The research question of the first article concerned the speed of lexical access in relation to L2 literacy. The conclusion was that speed of lexical access was found to increase with L2 proficiency, and it differentiated the age groups significantly. The gap between the speed of L1 and L2 lexical access also gradually decreased. The speed of lexical access managed to explain the variance in reading comprehension (20–30%) and essay writing (30–45%) quite substantially. In the lower grades (Grade 4 and the Primary school for the Russian cohort), word recognition speed explained the largest proportion of variance; in the higher grades (Grades 8 and 11), naming speed accounted for the largest proportion thereof. The contribution of speed of L2 word retrieval to L2 literacy is worth noticing, as this type of task is rarely used in L2 research or assessment. Furthermore, L1 measures also explained some of the variance in L2 literacy skills, and this was interpreted as a possible reflection of transfer of language-independent, cognitive skills.

The second article explored the role of accuracy in lexical access with a set of statistical analyses, controlling first for the speed measures. Accuracy in word recognition was shown to account for an independent proportion of variance in L2 reading comprehension and writing. The amounts were quite substantial (10–17%), and this connection was found across all grades. Of all the lexical access measures, accuracy in the L2 Word Chain task had the highest correlation with both literacy measures (except for Grade 8 writing), and it also accounted for the largest part of the variance. Therefore, it was concluded that in addition to the speed of lexical access, accuracy of lexical access is an important feature in literacy development, and a complimentary element in the fluency construct.

The third research article examined the inaccuracies from the perspective of inefficient processing and attention-control. The distributions of these were considered in relation to L2 proficiency levels. The first category, inefficient processing, was hypothesised to decrease with proficiency; whereas the second, attention-control, was proposed to stay stable or to increase. In the RAS L1 task, the largest part of the inaccuracies in all proficiency levels fell into the attention-control categories. In the RAS L2 task, in the youngest group, the inaccuracies were considered to have resulted from inefficiency, but in the higher grades, they were found to be more related to the attention-control categories. In the Word List task, perhaps in part due to its difficulty in L1 Finnish, both L1 and L2 inaccuracies were related to inefficiency (pronunciation problems). However, in the higher grades (8 and 11), the second most common type of inaccuracies in L1 were those relating to the use of attentional processes. Therefore, the distributions of inaccuracies were found to relate to proficiency to an extent, and this was considered to support the initial hypothesis of two differing trends in inaccuracies based on their background.

5 DISCUSSION

5.1 Theoretical considerations on cognitive fluency

The present study set out to explore both second and foreign language fluency, the main interest being in cognitive fluency and how it can serve as a reflection of the connections between fluency and proficiency. The background for this approach came from Segalowitz's (2000, 2010) fluency framework, and cognitive fluency was perceived as the efficiency and fluidity with which the cognitive processes underlying language use operate. This fluidity entails a balanced combination of automatic processes and attention-control mechanisms in speaking, reading, writing, and listening.

I suggested that cognitive fluency would be revealed in how fluently L2 users can access their mental lexicon. The first aspect of this is how efficiently lexical access works, that is, the speed and accuracy with which words are recognised or retrieved from long-term memory. Different level L2 learners, from two language backgrounds, were compared in their efficiency (speed and accuracy) of lexical access. The second aspect involves how capable a language learner is to use the resources available for attention-control and monitoring of his or her own production (see **Chapter 2.2**). The efficiency and attentional aspects of lexical access were examined in the form of inaccuracies that were proposed to be associated with these two sources. The inefficiency-related inaccuracies included guessing, skipping, within-category errors, and pronunciation difficulties. The attention-control inaccuracies included self-corrections, repetitions, and difficulties in inhibition and set-shifting. Furthermore, to distinguish the aspects of cognitive fluency from subjective speculations, performance in fluency of lexical access was compared to performance in reading comprehension and essay writing tasks in L2, to see to what extent fluency differences are able to reflect L2 proficiency. To find possible differences between the L2 specific aspects of fluency and the aspects of more general L1 fluency, similar lexical access tasks were conducted in both L1 and L2, and the relationships of those to

L2 literacy skills were examined. Next, the results from the research articles are considered in relation to the general research questions proposed in **Chapter 1.2**.

5.1.1 Aspects of cognitive fluency: efficiency and attention

Regarding the efficiency aspect of cognitive fluency, the speed of lexical access was found to increase with L2 proficiency, and the within-group variation also decreased as the students became faster (Article I). Both word recognition and word retrieval speed explained a significant amount of the L2 literacy skills. This is in accord with previous research in which these skills have been shown to relate to reading proficiency in both L1 and L2 (e.g. Dufva & Voeten, 1999; Kirby et al., 2003). This is especially interesting in the case of L2 writing, for which the previous results have been equivocal (Harrison et al., 2015; Schoonen et al., 2002). Therefore, the speed with which the mental lexicon is accessed was found to play an important part in language proficiency in this study. This finding should be further highlighted, as naming speed tasks have been rarely applied in L2 studies (apart from Gholamain & Geva, 1999).

The accuracy of lexical access was found to increase alongside L2 proficiency, and the children became more accurate in retrieving and recognising words (Article II). In the regression analyses, accuracy added more explained variance to the reading results than to the writing ones, which may relate to the tasks used. The Word Chain task results, which accounted for the most variance explained by accuracy, required some contextual considerations in addition to lexical access, as the texts were short stories in lieu of unrelated words. Therefore, it overlapped with the reading comprehension task to some extent (correlation r = .55). This contribution of accuracy must be underscored, however, as it is often overlooked but here contributed quite substantially to literacy scores. This means that there may be some trade-off between speed and accuracy, especially in speeded-up situations such as those of the tasks used here, in that a fast performance is not always an accurate one or vice versa. The Pearson correlations between the speed and accuracy in the tasks were, however, quite varied (for Word List L1 4: r=.542, 8: r=.182, and 11: r=.033; for List L2 4: r=.236, 8: r=.062, and 11: r=.168; for RAS L1 4: r=.144, 8: r=.213, and 11: r=.158; and for RAS L2 4: r=.260, 8: r=.372, and 11: r=.315). These correlations, therefore, serve to illustrate that these constructs measured somewhat different phenomena, although they are not enough to confirm or refute the trade-off interpretation. In L1 studies, usually only the speed of word retrieval is measured (cf. Bowey et al. 2004; Di Filippo et al. 2005; Kirby et al. 2003), and regarding L2 word recognition, sometimes it has been criticised that only accuracy has been accounted for (e.g. Geva et al., 2007). The overall interpretation of the current results, therefore, highlights how efficiency truly is a dual construct: both of its parts, speed and accuracy, seem to be important elements of more general fluency (cf. Jeon & Yamashita, 2014).

Examining and contrasting the inefficiency and attentional processes by means of the children's inaccuracies in lexical access revealed some interesting results and tendencies (Article III). Overall, the findings showed that proficiency levels influenced the types of inaccuracies that were produced. In the case of the L1 task versions and of the older students, the inaccuracies relating to attention allocation (especially self-corrections) accounted for the largest proportion of the total number of inaccuracies. On the other hand, in L2 tasks and for Grade 4, the difficulties seemed to arise more from the lack of efficiency (guessing, skipping, and pronunciation problems). The inaccuracies that were hypothesised to be related to inefficiency decreased with proficiency, and their share of the total amount of inaccuracies decreased as well. Pronunciation problems and guessings followed this trend most clearly. This pattern was as expected: with developing proficiency, the errors become fewer. Frequency played a role in the guessings in the Word List task, as the items often resulted in the more familiar words or forms (cf. Ellis, 2002; Gollan et al., 2011). Skipping patterns were not as clear, as the patterns differed somewhat in relation to language and proficiency. These differences were interpreted to have resulted from more monitoring resources being available in L1, and from strategic reasons, so this calls for caution in assessment of skippings, as they may also result from conscious choices (Kormos, 2000; Marian et al., 2013).

On the other hand, the inaccuracies connected with attention-control processes constituted the largest proportion of inaccuracies in the RAS L1 task, and this proportion increased in tandem with proficiency in the other tasks as well. Grade 4 was found to use self-corrections, an example of attention-control, significantly less than did the older students in both L1 and L2 tasks. This offers confirmation to the initial hypothesis that monitoring requires some proficiency level or threshold (see e.g. Gilabert, 2007; Lennon, 1990; Segalowitz, 2010). This may be related to 1) the development of language skills, which frees up more resources for higher aspects of language use; as well as to 2) the development of the whole monitoring system (see e.g. Diamond, 2013). Naturally, in addition to the lack of resources, beginning language learners also lack the knowledge to recognise erroneous forms and to provide an appropriate substitute. A more exact relationship of these two possible reasons should be examined further. Furthermore, the findings for repetitions, inhibition difficulties, and set-shifting suggest that the inaccuracies connected with the attentional process do not seem to be reliable indicators of proficiency (in line with e.g. Bosker et al., 2013; Freed, 2000). These results highlight the need for caution in the assessment; for example, self-corrections might be better interpreted along the lines of active control of resources instead of along those of breakdowns of spoken fluency. Repetitions were generally found to be very evenly distributed in all grades, the only exception being the RAS L1 task, where Grade 4 repeated slightly more often than the other groups. Here again we should consider the possibility of strategic repetition as a means of gaining processing time (cf. Engelhardt et al., 2010); overall, its use seems to depend on preferred speaking styles. Therefore, its inclusion either in the spoken fluency breakdowns (Skehan, 2003) or in the current attentional process category does not perhaps do it justice. Also, the effect of easier inhibition or lesser activation of lexical categories in L2 should be considered, although this would need further confirmation from tasks better suited to measuring selective attention and the allocation of it (see e.g. Favreau & Segalowitz, 1983; Jones et al., 2008; Marian et al., 2013).

Overall, considering the fluency model proposed in the theoretical chapter (**Chapter 2.2.2**; Segalowitz, 2000, 2003, 2010; see also Baddeley, 1996; Grabe, 2009; Just & Carpenter, 1992; Lennon, 2000; Miyake et al., 2000; Segalowitz & Frenkiel-Fishman, 2005; Skehan, 2003), the results from the research articles lend it support. Both speed and accuracy were found to contribute independently to the general fluency construct. The limitations in language skills, therefore, manifest themselves in inefficient lexical access, which gradually automatizes with rehearsing, and inefficiency decreases. When the low-level skills, such as lexical access, begin to operate automatically, they free up resources for the attention-requiring processes. These attentional processes, such as inhibition and set-switching, require a certain level of language skill. The differing tendencies found for the inefficient processing and the attention-control related phenomena are in line with the fluency model (cf. Kahng, 2014; Kormos, 2000; Marian et al., 2013; Samuels & Flor, 1997). The construct of executive functions was found to provide interesting viewpoint as a framework for examining attention-related inaccuracies, although more research is needed on its effects. Including at least the self-corrections in this category instead of repairs (Skehan's model, 2003) may offer a fruitful way to interpret the sometimes conflicting findings regarding them (e.g. Kahng, 2014; Kormos, 2000).

5.1.2 Fluent lexical access and connections to L2 proficiency

To control for the caveat that we measure only the subjective impressions of the researcher, the results from the fluent lexical access were compared to L2 proficiency, here measured with literacy tasks. The results of Articles I and II offer confirmation to the view that fluent lexical access is connected to reading comprehension and writing skills. Previously, lexical access tasks have been robustly connected to L1 reading development (Bowey et al., 2004; Di Filippo et al., 2005; Georgiou et al., 2013; Heikkilä et al., 2009; Kirby et al., 2003; Protopapas et al., 2013; Puolakanaho et al., 2007; Chapter 2.4.1), but the findings concerning L2 literacy have been equivocal regarding the extent of the interaction. That is why the current findings concerning L2 reading, and especially L2 writing, are highly interesting (and in agreement with, e.g. Perfetti's model, 2007). The results are in line with the theories of automatic sub-processes as precursors of literacy skills and proficiency.

In the current study, the lexical access tasks contributed to the variance in L2 literacy skills by 30–50%, which is quite substantial compared to many previous studies (e.g. Schoonen et al., 2002; Van Gelderen et al., 2007; Yamashita, 2013). Often wider measures, such as metacognition (Van Gelderen et al., 2007) or the size of vocabulary (cf. Joshi, 2005), have been shown to explain reading skills better than lexical access. Jeon and Yamashita (2014) showed in their meta-analysis that the shared variance between L2 comprehension and L2 vocabulary knowledge was 62%. Similar results were found in additional (unreported) analyses performed on the data available from the DIALUKI project, which showed that lexical access did not account for any significant additional variance in L2 literacy measures when L2 vocabulary scores were included. However, I suggest that the components influencing the building of a wide and componential skill, such as literacy, are equally interesting. Jeon and Yamashita (2014) included in the concept of vocabulary knowledge "a wide range of lexical knowledge dimensions, e.g., receptive knowledge, productive knowledge, vocabulary depth or breadth, knowledge of word form, meaning, or use" (p. 165). It is an interesting question of what this kind of much-encompassing construct really is informative.

Incidentally, in the aforementioned additional analyses on DIALUKI data, it was also found that the speed of lexical access accounted for around 40-50% of the variance in vocabulary scores. This is in line with De Jong and colleagues (2015), where L1 and L2 fluency explained 42% of vocabulary scores (see also, Jenkins et al., 2003). One interpretation for this might be that fluent reading is based on good vocabulary, which is in turn based on fluent lexical access (cf. Bialystok et al., 2008). Fluent lexical access could thus be a skill underlying both vocabulary learning and literacy development; the question of the direction of the influence is, however, still open for debate. As Jenkins and colleagues (2003) showed, list reading (lexical access), passage reading (vocabulary knowledge in context), and reading comprehension shared some of the variance with each other but not all, thus drawing partly from separate sub-processes of reading. Word recognition efficiency might, therefore, be understood as a baseline affecting how much, e.g., contextual cues can be used in comprehension (a higher-level process). The hypothesis in this thesis was also that lexical access taps into the more language-independent, cognitive language-learning processes, instead of language-dependent skills. Lexical access would in this way be one of the basic building blocks for an extensive (language-dependent) vocabulary, which then assists reading comprehension (this, in turn, helping the vocabulary acquisition, as Jenkins et al, 2003, point out; see also, Grabe, 2010). Further possibilities are, however, that the transparent nature of Finnish language affected results, as children are able to decode most words they encounter without necessarily knowing the meaning (therefore, their shared variance was only 50%). The substantial contribution of lexical access in the current study may, however, have been assisted by the combination of tasks: measuring word recognition and retrieval, using L1 and L2 tasks, and measuring separately speed and accuracy. This kind of approach perhaps made it possible to cover more of sub-processes of reading than a more limited set of tasks.

Further points worth noticing are that most previous studies have been conducted with university-level students (Akamatsu, 2008; Gollan et al., 2011; Yamashita, 2013). As the findings from Articles I and II highlighted that the skills underlying L2 reading and writing differed according to grade level, the influence of the L2 learners' age and proficiency on the results should be carefully considered. As word recognition (decoding) better explained the results of the younger children, and word retrieval (automatic access) did so for the older groups, this seems to reflect the gradual and possibly qualitative development of efficiency in lexical access (cf. Gollan et al., 2011). The different distributions of the inaccuracies in lexical access (Article III), based on the proficiency levels, offer further information concerning the developmental differences in lexical access.

5.1.3 The relationship of L1 and L2 fluency of lexical access

The need for studying the connections between L1 and L2 proficiency has been raised in certain studies (Alderson, 1984; Geva et al., 1997; Hulstijn, 2011), and this was possible with the current design. Regarding the correlations between the L1 and L2 versions, in the RAS task these increased in conjunction with the grade level (Article II). This may reflect how the task gradually begins to tap into the same, general cognitive speed as the L2 proficiency develops (the threshold as proposed by Alderson, 1984). The high correlations between the speed in the L1 and L2 Word Lists suggest they reflected the same process as well, likely related to the use of decoding skills of alphabetic languages. On the otherhand, the lower correlations between the accuracy scores in both tasks, as compared to the speed results, may reflect that accuracy is a more language-dependent skill than is speed, although there was such little variance in the RAS task that this may have affected the results as well. These results echo, however, what Van Gelderen and colleagues (2007) have proposed, that L1 and L2 reading begin to use the same skills as proficiency develops. This is reflected also in the results of Articles I and II, that L1 fluency had a bearing on the L2 literacy, even in the higher grades. For example, the speed in the RAS L1 task had the highest correlation to the literacy measures in Grade 8, where it also contributed to the explained variance in literacy measures significantly. This might be interpreted as an example of the required level of L2 skills for the L1 influence to appear (see **Chapter 2.5**; Alderson, 1984; Bernhardt, 2005; Hulstijn, 2011; Koda, 2008). However, the accuracy in the L1 Word List task seemed most related to the literacy measures in Grade 4. It also correlated with the literacy measures even more than accuracy in the L2 Word List task, which was not the case for the older groups (see Article II). This may be interpreted as reflecting a phase when one has studied the L2 only a little over one year, and the decoding processes even in L1 are still not automatized (in line with Cheung & Lin, 2005). Therefore, no evidence regarding the thresholds of L1 skills in lexical access (Cummins, 1991) was found in the current study, but some evidence pointing to the L2 thresholds (Alderson, 1984) may be seen. The results may be interpreted, however, in such a way that the reason behind the transfer may be different in the different phases of L2 learning, reflecting the differences between the language-dependent and language-independent processes.

Another important point is that the current comparisons revealed insights into the disfluent nature of L1, as was illustrated by the sheer number of inaccuracies in L1 (Article III). This confirms the need for caution in the indiscriminate L2 fluency assessments. As has been shown previously by Kormos (2000), the types of inaccuracies, however, differed between the languages. Therefore, research into the different types and sources behind the inaccuracies is important.

Furthermore, it seems warranted to control for L1 performance to distinguish personal speaking styles from proficiency (e.g. De Jong et al., 2015). The use of repetitions and self-corrections has been previously proposed as being more related to language strategies, which are developed in L1 and perhaps also transferred as a strategy for buying time (Kormos, 1999; Peltonen & Lintunen, 2016). Therefore, they may require a threshold of proficiency before being applied in L2, and this was the tendency found in the current study (Article III). Lehtonen (1978: 67) has pointed out: "To be fluent [..] one has to know how to hesitate, be silent, self-correct, interrupt [..] in a way that is expected in the linguistic community". Consequently, the question is transformed as to what extent should a speaker be compared to an idealised native speech community, and to what extent should his or her own speaking style be considered? This question is no closer to being answered.

5.1.4 Lexical access and its role in cognitive fluency

As the results of the articles demonstrated, the respective contributions of word retrieval and word recognition were different for different-age learners, as well as for second and foreign language learners (see **Chapter 4.1**). This offers confirmation to the proposition that these tasks measure somewhat different aspects of lexical access, although in a complimentary manner (cf. Georgiou et al., 2013; Gholamain & Geva, 1999; Gollan et al., 2011; Snellings et al., 2002; **Chapter 2.2**). Using both measures may thus reveal a more complete picture than either one alone, and, contrary to what Jenkins et al. (2003) claim, a word list cannot be used to measure naming speed. Word recognition tasks measure the fluency of decoding and sight-word recognition; word retrieval (especially when operationalised as naming speed) is reflective of the automaticity of very familiar vocabulary. The results in the current thesis may be interpreted as reflecting the change in the relationship of these skills during the language learning process. This is supported by the different skill pattern between the second and foreign language learners, as the second language learners' L1 skills were at the same level as that of their same-age peers (primary school), but their L2 skills were closer to those of the secondary school students. More research and further analyses are naturally needed to corroborate these conclusions, but they are one possibility for explaining and interpreting the differing results between different learner groups. As such, these results bear a resemblance to what Harrison and colleagues (2015) report, that the writing achievement of English L1 and L2 speakers was partly explained by different predictors. According to their results, the measures of naming speed had more influence on the spelling and writing quality for the L2 writers than for the L1 ones.

The proposition that fluent lexical access would have more bearing at the beginning of L2 learning was, however, not corroborated (cf. Durgunoğlu et al., 1993; Jenkins et al., 2003). The explained variance in fact increased with proficiency, and was highest for the second language learners. One explanation for this may have to do with the nature of the tasks: the RAS task, which accounted for the most variance for the older groups, may tap further into the more general processing fluency than the word recognition tasks most often used in L2 studies, as was highlighted in the previous chapter. If the tasks can reflect the general processing capacity, as the RAS one seems to do, it offers further justification for the approach and employment of lexical access for measuring more than just articulation or surface fluency features. As the data acquired from lexical access tasks is more controlled than from speech elicitation tasks, it is also less dependent on the conversational or contextual requirements. The current results of the relationship of fluent lexical access and L2 reading and writing skills are in line with this notion.

5.2 Limitations and future directions

Next, the limitations of the current study and possible remedies to them are discussed. Several issues emerged with the task constructs (which can also be viewed as interesting findings). Most of the tasks used derive from L1 research, and their application in L2 testing contributes important insights. The comparability of the tasks was an issue with the L2 reading tasks and with the different language versions. Concerning the L2 reading tasks, both the content-knowledge and proficiency are widely different for children along this age spectrum. Furthermore, the comparability between the age groups was sometimes challenging, as the same reading comprehension tasks could obviously not be used. Reading skills are both quantitatively and qualitatively different in primary school compared to upper secondary school. The comparability was thus difficult to ascertain, when the tasks could not be used to measure exactly the same reading sub-skills. The different language versions were made to be as comparable as possible, but this was a difficult task. For the Word List, the versions contained an equal number of words but differed greatly in the number of syllables. On the other hand, the Word List was a good example of a task that was the same for all age groups, and being challenging enough, it distinguished well between all the students. Nevertheless, using the same task for all groups was problematic in some instances, because this could result in too-easy tasks for the older students. This was the case especially with the Rapidly Presented Words (RPW) task, where there was a ceiling effect, the consequence being that it could not be included in the later analyses. Nevertheless, the task showed some promise in Grade 4 in additional analyses. This raises the interesting question of whether a more challenging version could also be used for discerning older students, as it measures the whole-word reading that develops gradually with automaticity. It would be fascinating, furthermore, to determine if there are differences in the task between different orthographies, as for example the syllabic nature of Japanese kana script has been proposed to make whole-word recognition in an alphabetic L2 more challenging (Akamatsu, 2008).

For the Word Chain task, measuring the speed more accurately (at shorter intervals) than to the nearest 30 seconds probably would have yielded more accurate results. This was unfortunately not possible in the large group settings.

On the other hand, this, along with the finding that Word Chain is not a simple lexical access task, can be considered as useful findings pertaining to future application of these tasks. Thus, if the aim is to measure lexical access more 'purely', a Word Chain task with unconnected words might be used (cf. Akamatsu, 2008). If the aim is to gather as much information on someone's reading skills efficiently, the contextual excerpts seem to serve better. Furthermore, as the current study was a cross-sectional one, a longitudinal setting would definitely be needed to confirm the results concerning the changes in fluency profiles within the individuals (cf. Riggenbach, 1993).

The heterogeneity of the Russian results was a drawback, as they could not be included in Articles II and III. This was particularly disappointing because the results from the first paper suggested that comparing second and foreign language learners might reveal highly interesting differences. The development of second language learners would be of most interest, as many of the primary school-level children are either learning to read simultaneously in two languages and alphabets, or only in a less proficient language. The literacy development of second language learners is very important in every society striving for inclusion and integration of immigrants, and this requires knowledge of how they differ as learners from L1 learners (for discussion on L1 and L2 reading development differences, see e.g. Alderson, 1984; Gholamain & Geva, 1999).

In discussing the analysis of accuracy, I highlighted that the categories and results were quite tentative. The skewness in their number was a problem for the statistical analyses. The reasons behind the inaccuracies were in part hypothesised only indirectly; nevertheless, the interrater agreement was at an acceptable level. More solid confirmation is, however, naturally needed. These results a best interpreted as a starting point for further studies, especially in the case of inhibition and set-shifting difficulties, whose relationships to fluent lexical access should be explored with tasks designed for these phenomena (see e.g. Bialystok et al., 2008; McCabe et al., 2010; Miyake et al., 2000). More direct studies on the relationship of voluntary and accidental skippings would also be highly interesting, for example with the incorporation of an eye-tracking paradigm.

5.3 Concluding remarks

Hopefully, I have managed to illustrate throughout this thesis that fluency is a more varied phenomenon than it might at first appear in, for example, assessment criteria. The results do lend some support to the use of lexical access tasks as quick and easily administered tools for assessing L2 proficiency, as the fluency of word recognition and retrieval were shown to explain variance in L2 reading and L2 writing quite well. This was, however, the case only when the tasks measured aspects of inefficient lexical access. When the inaccuracies reflect more the utilisation of attentional processes, caution in the interpretation of the results of the tasks should be exercised. As the tasks stem from diagnostic use, perhaps the most beneficial application of them would still be for purposes of screening for possible problems rather than for those of overall assessment. Furthermore, the disfluencies in L1 use must be highlighted, and an interesting question remains: how can we describe and distinguish native-like disfluency? Overall, as any study is wont to do, this one has raised more questions than has answered, but I have tried to demonstrate that combining theories and practices from different disciplines and frameworks can provide novel ways of looking at a chosen construct. This approach may be helpful, for example, in validating and broadening the scope of the tasks and the methods with which they are usually assigned. Especially when one examines a phenomenon as slippery and moving as fluency, this multi-disciplinary approach is proposed as a fruitful way of tapping into the different aspects of fluency. The aim has been to underline that strict, inflexible definitions hardly do justice in any way, shape, or form to the variability of speakers, and to their command and use of languages.

YHTEENVETO

Toisen ja vieraan kielen sujuvuus kognitiivisesta näkökulmasta: sananhaun tehokkuus ja tarkkaavaisuuden säätely

Vieraan kielen puhujia arvioidaan jatkuvasti sen mukaan, kuinka sujuvaa heidän kielenkäyttönsä on. Myös kielenopetuksessa ja virallisissa arviointiohjeissa sujuvuus on yksi taitokriteereistä (mm. EVK, Council of Europe, 2007; POPS, 2014[6]). Tämä on ongelmallista, sillä sujuvuuteen vaikuttavat monet seikat, kuten persoonallinen puhetyyli. Tästä syystä sujuvuustutkimus onkin keskittynyt etsimään puheesta sellaisia mitattavia ominaisuuksia, jotka luotettavimmin kertoisivat kielitaidon tasosta (De Jong ym., 2015). Yhteisymmärrystä parhaista ennustajista ei ole kuitenkaan saavutettu.

Väitöstutkimuksessani tarkastelen toisen ja vieraan kielen sujuvuutta kognitiivisesta näkökulmasta. Väitöskirjani koostuu kolmesta artikkelista ja yhteenvedosta. Teoreettisena taustana on Segalowitzin (2000, 2010) teoria, jossa sujuvuus jaetaan kolmeen osa-alueeseen: kognitiiviseen sujuvuuteen, puheen sujuvuuteen ja havaittuun sujuvuuteen. Esitän väitöskirjassani, että jos sujuvuutta halutaan käyttää arvioinnin välineenä, olisi ensin ymmärrettävä, miten puheen pinnalliset sujuvuusilmiöt suhteutuvat kognitiiviseen sujuvuuteen. Tutkimuskysymykseni käsittelevät toisen ja vieraan kielen sujuvuuden mittaamista, sujuvuuden suhdetta toisen ja vieraan kielen luku- ja kirjoitustaitoihin sekä suhdetta äidinkielen sujuvuuteen.

Lähtöoletuksena ovat teoriat kognitiivisten prosessien rajallisuudesta (esim. Just & Carpenter, 1992; Kormos, 1999; Grabe, 2009). Kognitiiviseen sujuvuuteen sisältyy kaksi erisuuntaista prosessia: taitojen automaattistuminen sekä tarkkaavaisuuden säätely (Segalowitz, 2000). Mitä automaattisemmin alemman tason prosessit, kuten sananhaku, toimivat, sitä enemmän rajallisia resursseja jää tarkkaavaisuuden säätelyyn. Kognitiivisen sujuvuuden mittaamisen työkaluksi esitän sananhaun sujuvuutta. Sujuvan sananhaun on nimittäin osoitettu ennustavan myöhemmin kehittyvää lukivaikeutta jo hyvin varhaisessa vaiheessa (esim. Puolakanaho ym., 2007).

Tehokas lukeminen vaatii sujuvaa sanantunnistusta (Ehri, 1991) ja kirjoittaminen sujuvaa sananhakua (McCutchen, 2000). Sujuvuus voidaan määritellä koostuvaksi sekä nopeudesta ja että tarkkuudesta, joten mittasin näitä kumpaakin sananhaussa (sanalistan lukeminen, nopeasti välähtävät sanat, nopean nimeämisen testi RAS) sekä äidinkielellä että toisella tai vieraalla kielellä. Näitä tuloksia verrattiin toisen ja vieraan kielen luetunymmärtämisen ja kirjoittamisen tuloksiin. Koehenkilöinä (yhteensä N=821) oli kolme ikäryhmää suomalaislapsia (4., 8. ja lukion toiselta luokalta), jotka opiskelivat englantia vieraana kielenä. Toisena ryhmänä testattiin alakouluikäisiä venäläistaustaisia lapsia, jotka opiskelivat suomea toisena kielenä. Tämä asetelma mahdollisti vertailun sekä eri-ikäisten että eri kielitaustasta tulevien oppilaiden välillä.

Artikkelissa I tutkin regressioanalyysin avulla kuinka paljon nimeämis- ja sanantunnistuksen nopeus selittivät toisen kielen tekstitaitoja. Tuloksissani sananhaun nopeus selitti toisen ja vieraan kielen luetunymmärtämisen ja kirjoittamisen tuloksissa esiintynyttä vaihtelua merkittävässä määrin (20–45%). Eri ikäryhmien eroista merkittävin löydös oli, että *sanantunnistuksen* nopeus selitti nuorempien oppilaiden tuloksia, kun taas vanhempien opiskelijoiden taitoja selitti enemmän *nimeämisnopeus*. Myös äidinkielen sanahaun sujuvuus selitti vieraan kielen tekstitaitojen vaihtelua, vaikka selitysosuus jäi pienehköksi. Artikkelissa II tarkastelin, lisääntyikö luku- ja kirjoitustaidoissa esiintyvän variaation selitysaste, kun selittäväksi tekijäksi otettiin mukaan myös sananhaun tarkkuus. Tarkkuuden huomioon ottaminen lisäsi selitysosuutta 10–17%. Tämä on merkittävä tulos, sillä usein vain sananhaun nopeus on ollut tutkimuksen fokuksessa.

Artikkeli III oli lähestymistavaltaan laadullisempi. Tarkastelin siinä sananhaun tarkkuutta sanalistassa ja nopean nimeämisen tehtävässä (RAS) eri ikäryhmien välillä. Tarkkuutta mittasin jaottelemalla epäsujuvuudet niin, että ne johtuivat joko automaattistumisen puutteesta tai tarkkaavaisuuteen käytettävien resurssien lisääntymisestä. Hypoteesina oli, että osa epäsujuvuuksista voi esiintyä vasta kun kielitaitoa on riittävästi ylemmän tason prosessien käyttöön (esim. Gilabert, 2007; De Jong ym., 2015). Esimerkiksi itsekorjauksia ajateltiin esiintyvän vähemmän vieraan kielen alemmilla taitotasoilla (Kormos, 2000; Jarvis ym., 2014), koska niiden tekemiseen ei ole vielä resursseja eikä taitoja. Englantia vieraana kielenä opiskelevien suomenkielisten lasten tekemien virheiden määrät ja suhteelliset osuudet laskettiin sanalistatehtävässä ja nimeämistehtävässä (RAS) ja niitä vertailtiin luokka-asteittain. Analyysien tulokset osoittivat, että automaattisuuden puutteeseen liittyvät virheet vähenivät kielitaidon kasvaessa ja niitä esiintyi enemmän vieraalla kuin äidinkielellä. Tarkkaavuuteen liittyvät virheet sen sijaan lisääntyivät 4. luokalta 8. luokalle, mutta vähenivät taas lukiolaisten tuloksissa. Lisäksi, varsinkin nuoremmilla oppilailla, tarkkaavuuteen liittyviä epäsujuvuuksia esiintyi enemmän äidin- kuin vieraalla kielellä.

Tutkimuksen tulokset antavat tukea teoreettiselle mallille, jonka mukaan kognitiivisen sujuvuuden erisuuntaiset prosessit, automaattistuminen ja tarkkaavuus, heijastuvat myös eri tavoin suullisessa sujuvuudessa. Tarkkaavaisuuteen liittyvät epäsujuvuudet eivät näyttäneet tulosten valossa palvelevan suoraviivaista arviointikäyttöä. Väitöstutkimuksessani saadut tulokset vahvistavat näkemystä, jonka mukaan kognitiivista sujuvuutta on mahdollista mitata sanantunnistuksen ja -haun avulla. Lisäksi käytetyt mittarit selittivät itsenäisiä osia tekstitaidoista ja selitysosuudet erosivat ikäryhmien välillä, joten sananhaun eri puolia mittaavien testien yhdistämisestä näyttäisi olevan hyötyä. Kai-ken kaikkiaan sujuvuus näyttäytyy työn valossa niin moniuloitteisena ilmiönä, että sen tutkimuksen ja arvioinnin soisi heijastelevan tätä sortumatta liian jäyk-kiin määritelmiin.

Avainsanat: toinen ja vieras kieli, kognitiivinen sujuvuus, sanantunnistus, sanahaku, luku- ja kirjoitustaidot, virheanalyysi

REFERENCES

Ahonen, T., Tuovinen, S. & Leppäsaari, T. 2003. *Nopean sarjallisen nimeämisen testi*. [Test for Rapid Alternating Stimuli.] Niilo Mäki Institute.

Akamatsu, N. 2008. The effects of training on automatization of word recognition in English as a foreign language. *Applied Psycholinguistics*, 29, 175–193.

Alanen, R., Huhta, A., Jarvis, S., Martin, M., & Tarnanen, M. 2012. Issues and challenges in combining SLA research and language testing. In D. Tsagari & I. Csépes (Eds.), *Collaboration in language testing and assessment. Language Testing and Evaluation Series* (pp. 15–30). Frankfurt am Main: Peter Lang.

Alderson, J. C. 1984. Reading in a foreign language: A reading problem or a language problem? In J. C. Alderson, & A. H. Urquhart (Eds.), *Reading in a foreign language* (pp. 1–27). Harlow: Longman.

Alderson, J. C., & Huhta, A. 2005. The development of a suite of computer-based diagnostic tests based on the Common European Framework. *Language Testing*, 22, 301–320.

Aro, M., & Wimmer, H. 2003. Learning to read: English in comparison to six more regular orthographies. *Applied Psycholinguistics*, 24, 621–635.

Baddeley, A. 1996. Exploring the Central Executive. *The Quarterly Journal of Experimental Psychology Section A*, 49, 5–28.

Balota, D. A., Yap, M. J., & Cortese, M. J. 2006. Visual word recognition: The journey from features to meaning (A travel update). In M. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 285–375). NY: Elsevier.

Bernhardt, E. 2005. Progress and procrastination in second language reading. *Annual Review of Applied Linguistics*, 25, 133–150.

Bialystok, E., Craik, F. I. M., & Luk, G. 2008. Lexical access in bilinguals: Effects of vocabulary size and executive control. *Journal of Neurolinguistics*, 21, 522–538.

Booth, J. R., Perfetti, C. A., & McWhinney, B. 1999. Quick, automatic and general activation of orthographic and phonological representations in young readers. *Developmental Psychology*, 35, 3–19.

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

Bowey, J. A., Storey, T., & Ferguson, A. N. 2004. The association between continuous naming speed and word reading skill in fourth- to sixth-grade children. *Australian Journal of Psychology*, 56, 155–163.

Broadbent, D. E. 1967. Word-frequency effect and response bias. *Psychological Review*, 74, 1–15.

Cattell, J. M. 1886. The time it takes to see and name objects. *Mind*, 11, 63–65.

Catts, H., Gillispie, M., Leonard, L. B., Kail, R., & Miller, C. A. 2002. The role of speed of processing, rapid naming, and phonological awareness in reading achievement. *Journal of Learning Disabilities*, 35, 509–524.

Cheung, H., & Lin, A. M. Y. 2005. Differentiating between automatic and strategic control processes: Toward a model of cognitive mobilization in bilingual reading. *Psychologia*, 48, 39–53.

Cook, V. 2002. *Portraits of the L2 user*. Buffalo, NY: Multilingual matters.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. 2001. DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108, 204–256.

Council of Europe. 2007. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.

Cummins J. 1991. Interdependence of first- and second-language proficiency in bilingual children. In E. Bialystok (ed.), *Language processing in bilingual children* (pp. 70–89). Cambridge: Cambridge University Press.

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

De Jong, P. F. 2011. What discrete and serial Rapid Automatized Naming can reveal about reading. *Scientific Studies of Reading*, 15, 314–337.

Denckla, M. B., & Cutting, L. E. 1999. History and significance of Rapid Automatized Naming. *Annals of Dyslexia*, 49, 29–42.

Denckla, M. B., & Rudel, R. G. 1976. Rapid "automatized" naming (R.A.N.): Dyslexia differentiated from other learning disabilities. *Neuropsychologia*, 14, 471–479.

Diamond, A. 2013. Executive functions. *Annual Review of Psychology*, 64, 135–68.

Di Filippo, G., Brizzolara, D., Chilosi, A., De Luca, M., Judica, A., Pecini, C., Spinelli, D., & Zoccolotti, P. 2005. Rapid naming, not cancellation speed or articulation rate, predicts reading in an orthographically regular language (Italian). *Child Neuropsychology*, 11, 349–361.

Drijbooms, E., Groen, M. A., & Verhoeven, L. 2015. The contribution of executive functions to narrative writing in fourth grade children. *Reading & Writing: An Interdisciplinary Journal*, 29, 989–1011.

Dufva, M., & Voeten, M. J. M. 1999. Native language literacy and phonological memory as prerequisites for learning English as a foreign language. *Applied Psycholinguistics*, 20, 329–348.

Durgunoğlu, A., Nagy, W., & Hancin-Bhatt, B. 1993. Cross-language transfer of phonological awareness. *Journal of Educational Psychology*, 85, 453–465.

Ehri, L. 1991. Development of the ability to read words. In R. Barr, M. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (pp. 383–417). New York: Longman.

Ellis, N. C. 2002. Frequency effects in language acquisition: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24, 143–188.

Engelhardt, P. E., Corley, M., Nigg, J. T., & Ferreira, F. 2010. The role of inhibition in the production of disfluencies. *Journal Memory & Cognition*, 38, 617–628.

Freed, B. F. 2000. Is fluency, like beauty, in the eyes (and ears) of the beholder? In H. Riggenbach (ed.), *Perspectives on fluency* (pp. 243–265). Ann Arbor, MI: University of Michigan.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. 2001. Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, 5, 239–256.

Georgiou, G. K., Papadopoulos, T. C., & Kaizer, E. L. 2014. Different RAN components relate to reading at different points in time. *Reading & Writing: An Interdisciplinary Journal*, 27, 1379–1394.

Georgiou, G. K., Parrila, R., Cui, Y., & Papadopoulos, T. C. 2013. Why is rapid automatized naming related to reading? *Journal of Experimental Child Psychology*, 115, 218–225.

Gernsbacher, M. A. 1993. Less skilled readers have less efficient suppression mechanisms. *Psychological Science*, 4, 294–298.

Geva, E., & Massey-Garrison, A. 2013. A comparison of the language skills of ELLs and monolinguals who are poor decoders, poor comprehenders, or normal readers. *Journal of Learning Disabilities*, 46, 387–401.

Geva, E., & Ryan, E. B. 1993. Linguistic and cognitive correlates of academic skills in first and second languages. *Language Learning*, 43, 5–42.

Geva, E., & Siegel, L. S. 2000. The role of orthography and cognitive factors in the concurrent development of basic reading skills in bilingual children. *Reading and Writing: An Interdisciplinary Journal*, 12, 1–30.

Geva, E., Wade-Woolley, L., & Shany, M. 1997. Development of reading efficiency in first and second language. *Scientific Studies of Reading*, 1, 119–144.

Gholamain, M., & Geva, E. 1999. Orthographic and cognitive factors in the concurrent development of basic reading skills in English and Persian. *Language Learning*, 49, 183–217.

Gilabert, R. 2007. Effects of manipulating task complexity on self-repairs during L2 oral production. *International Review of Applied Linguistics in Language Teaching (IRAL)*, 45, 215–240.

Gollan, T. H., Slattery, T. J., Goldenberg, D., van Assche, E., Duyck, W., & Rayner, K. 2011. Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, 140, 186–209.

Grabe, W. 2009. *Reading in a second language: Moving from theory to practice*. New York: Cambridge University Press. 2009.

Grabe, W. 2010. Fluency in reading – thirty-five years later. *Reading in a Foreign Language*, 22(1), 71–83.

Harrison, G. L., Goegan, L. D., Jalbert, R., McManus, K., Sinclair, K., & Spurling, J. 2015. Predictors of spelling and writing skills in first- and second-language learners. *Reading & Writing: An Interdisciplinary Journal*, 29, 69–89.

Hayes, J. R. 1996. A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–27). Mahwah, NJ: Lawrence Erlbaum.

Heikkilä, R., Närhi, V., Aro, M., & Ahonen, T. 2009. Rapid automatized naming and learning disabilities: Does RAN have a specific connection to reading or not? *Child Neuropsychology*, 15, 343–358.

Holopainen, L., Kairaluoma, L., Nevala, J., Ahonen, T., & Aro, M. (Eds.). 2004. *Lukivaikeuksien seulontamenetelmä nuorille ja aikuisille*. [Dyslexia screening test for youth and adults.] Jyväskylä. Niilo Mäki Instituutti.

Housen, A., & Kuiken, F. 2009. Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30, 461–473.

Hulstijn, J. H. 2011. Explanations of associations between L1 and L2 literacy skills. In M.S. Schmid & W. Lowie (Eds.), *Modeling bilingualism: From structure to chaos* (pp. 85–111). Amsterdam: Benjamins.

Hulstijn, J., Van Gelderen, A., & Schoonen, R. 2009. Automatization in second-language acquisition: What does the coefficient of variation tell us? *Applied Psycholinguistics*, 30, 555–582.

Häyrinen, T., Serenius-Sirve, S., & Korkman, M. 1999. *Lukilasse. Lukemisen, kirjoittamisen ja laskemisen seulontatestistö peruskoulun ala-asteen luokille 1–6*. [The Lukilasse graded achievement package for comprehensive school age children.] Helsinki: Psykologien Kustannus Oy.

Jarvis, S., Casal, J. E., Maloney, J., Taylor, J., Amira, M., Hill, J. M., Konstantinova, I., & Wistner, B. 2014. *Sentence memory and sentence processing: A longitudinal study of the effects of crosslinguistic similarity and learners' cognitive capacities*. Paper presented at the annual meeting of the American Association for Applied Linguistics (AAAL) 2014, Portland, OR, March 22–25.

Jenkins, J. R., Fuchs, L. S., van den Broek, P., Espin, C., & Deno, S. L. 2003. Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology*, 95, 719–729.

Jeon, E. H. 2012. Oral reading fluency in second language reading. *Reading in a Foreign Language*, 24, 186–208

Jeon, E.H., & Yamashita, J. 2014. L2 reading comprehension and its correlates: A meta-analysis. *Language Learning*, 64, 160–212.

Jones, M. W., Obregón, M., Kelly, M. L., & Branigan, H. P. 2008. Elucidating the component processes involved in dyslexic and non-dyslexic reading fluency: An eye-tracking study. *Cognition*, 109, 389–407.

Joshi, R. M. 2005. Vocabulary: A critical component of comprehension. *Reading and Writing Quarterly*, 21, 209–219.

Just, M. A., & Carpenter, P. A. 1992. A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122–149.

Kaakinen, J. K., Hyönä, J., & Keenan, J. M. 2002. Perspective effects on on-line text processing. *Discourse Processes*, 33, 159–173.

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

Kail, R., & Hall, L. K. 1994. Processing speed, naming speed, and reading. *Developmental Psychology*, 30, 949–954.

Kail, R., Hall, L. K., & Caskey, B. J. 1999. Processing speed, exposure to print, and naming speed. *Applied Psycholinguistics* 20, 303–314.

Kame'enui, E., & Simmons, D. 2001. Introduction to this special issue: The DNA of reading fluency. *Scientific Studies of Reading*, 5, 203–210.

Kieffer, M. J., Vukovic, R. K., & Berry, D. 2013. Roles of attention shifting and inhibitory control in fourth-grade reading comprehension. *Reading Research Quarterly*, 48, 333–348.

Kirby, J. R., Parrila, R. K., & Pfeiffer, S. L. 2003. Naming speed and phonological awareness as predictors of reading development. *Journal of Educational Psychology*, 95, 453–464.

Koda, K. 1996. L2 word recognition research: A critical review. *The Modern Language Journal*, 80, 450–460.

Koda, K. 2008. Impacts of prior literacy experience on second language learning to read. In K. Koda & A. M. Zehler (Eds.), *Learning to read across languages: Cross-linguistic relationships in first- and second-language literacy development* (pp. 68–93). London: Routledge.

Koponen, M., & Riggenbach, H. 2000. Overview: Varying perspectives on fluency. In H. Riggenbach (ed.), *Perspectives on fluency* (pp. 5–22). Ann Arbor, MI: University of Michigan.

Kormos, J. 1999. The effect of speaker variables on the self-correction behaviour of L2 learners. *System*, 27, 207–221.

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

Kroll, J., & Sunderman, G. 2003. Cognitive processes in L2 learners and bilinguals. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 104–129). Malden, MA: Blackwell.

Kuiken, F., Vedder, I., & Gilabert, R. 2010. Communicative adequacy and linguistic complexity in L2 writing. In I. Bartning, M. Martin & I. Vedder (Eds.), *Communicative proficiency and linguistic development: Intersections between SLA and language testing research* (pp. 81–100). Eurosla Monographs 1. Roma: Eurosla.

LaBerge, D., & Samuels, S. J. 1974. Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6, 293–323.

Lehtonen, J. 1978. On the problems of measuring fluency. In M. Leiwo & A. Räsänen (Eds.), *AFinLA yearbook 1978* (pp. 53–68). Jyväskylä: AFinLA.

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

Lennon, P. 2000. The lexical element in spoken second language fluency. In H. Riggenbach (ed.), *Perspectives on fluency* (pp. 25–42). Ann Arbor, MI: University of Michigan.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–38.

Logan, J. A. R., Schatschneider, C., & Wagner, R. K. 2011. Rapid serial naming and reading ability: the role of lexical access. *Reading & Writing: An Interdisciplinary Journal*, 24, 1–25.

Lyytinen, H., Eklund, K., Erskine, J., Guttorm, T.K., Laakso, M.-L., Leppänen, P.H.T., Lyytinen, P., Poikkeus, A.-M., Richardson, U., & Torppa, M. 2004. The development of children at familial risk for dyslexia: birth to school age. *Annals of Dyslexia*, 54, 185–220.

MacLeod, C. 1991. Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109, 163–203.

Manchón, R. M., Murphy, L., & Roca de Larios, J. 2007. Lexical retrieval processes and strategies in second language writing: A synthesis of empirical research. *International Journal of English Studies*, 7, 149–174.

Marian, V., Blumenfeld, H. K., Mizrahi, E., Kania, U., & Cordes, A.-K. 2013. Multilingual Stroop performance: Effects of trilingualism and proficiency on inhibitory control. *International Journal of Multilingualism*, 10, 82–104.

McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. 2010. The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology*, 24, 222–243.

McCutchen, D. 2000. Knowledge, processing, and working memory: Implications for a theory of writing. *Educational Psychologist*, 35, 13–23.

McVay, J. C., & Kane, M. J. 2012. Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology: General*, 141, 302–320.

Meuter, R. F. I., & Allport, A. 1999. Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, 40, 25–40.

Miller-Guron, L., & Lundberg, I. 2000. Dyslexia and second language reading: A second bite at the apple? *Reading & Writing: An Interdisciplinary Journal*, 12, 41–61.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. 2000. The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49–100.

Nation, K., & Snowling, M. J. 2004. Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, 27, 342–356.

Närhi, V., Ahonen, T. Aro, M., Leppäsaari, T., Korhonen, T. T., Tolvanen, A., & Lyytinen, H. 2005. Rapid serial naming: Relations between different stimuli and neuropsychological factors. *Brain and Language*, 92, 45–57.

Peltonen, P., & Lintunen, P. Published online January 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*.

Perfetti, C. A. 1985. *Reading ability*. New York: Oxford University Press.

Perfetti, C. A. 2007. Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11, 357–383.

Pollatsek, A., Rayner, K., & Balota, D. A. 1986. Inferences about eye movement control from the perceptual span in reading. *Perception & Psychophysics*, 40, 123–130.

Protopapas, A., Fakou, A., Drakopoulou, S., Skaloumbakas, C., & Mouzaki, A. 2013. What do spelling errors tell us? Classification and analysis of spelling errors of Greek schoolchildren with and without dyslexia. *Reading & Writing: An Interdisciplinary Journal*, 26, 615–646.

Puolakanaho, A., Ahonen, T., Aro, M., Eklund, K., Leppänen, P., Poikkeus, A-M., Tolvanen, A., Torppa, M., & Lyytinen, H. 2007. Very early phonological and language skills: Estimating individual risk of reading disability. *Journal of Child Psychology and Psychiatry*, 48, 923–931.

Salmi, P. 2008. *Nimeäminen ja lukemisvaikeus. Kehityksen ja kuntoutuksen näkökulma*. [Naming and Dyslexia: Developmental and Training Perspectives.] Doctoral dissertation. Jyväskylä Studies in Education, Psychology and Social Research, 345.

Samuels, S. J., & Flor, R. F. 1997. The importance of automaticity for developing expertise in reading. *Reading & Writing Quarterly*, 13, 107–121.

Sasaki, M. 2002. Building an empirically-based model of EFL learners' writing processes. In S. Ransdell & M-L. Barbier (Eds.), *New directions for research in L2 writing* (pp. 49–80). Dordrecht: Kluwer.

Schmidt, R. 1992. Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, 14, 357–385.

Schoonen, R., Van Gelderen, A., De Glopper, K., Hulstijn, J., Snellings, P., Simis, A., & Stevenson, M. 2002. Linguistic knowledge, metacognitive knowledge and retrieval speed in L1, L2 and EFL writing. A structural equation modelling approach. In S. Ransdell & M.-L. Barbier (Eds.), *New Directions for Research in L2 Writing* (pp. 101–122). Dordrecht: Kluwer.

Seidenberg, M. S., & McClelland, J. L. 1989. A distributed developmental model of word recognition and naming. *Psychological Review*, 96, 523–568.

Segalowitz, N. 2000. Automaticity and attentional skill in fluent performance. In H. Riggenbach (ed.), *Perspectives on fluency* (pp. 200–219). Ann Arbor, MI: University of Michigan Press.

Segalowitz, N. 2003. Automaticity and second languages. In C. J. Doughty & M. H. Long (Eds.), *Handbook of Second Language Acquisition* (pp. 382–408). Malden, MA: Blackwell Publishing.

Segalowitz, N. 2010. *Cognitive bases of second language fluency*. New York: Routledge.

Segalowitz, N., & Frenkiel-Fishman, S. 2005. Attention control and ability level in a complex cognitive skill: Attention-shifting and second language proficiency. *Memory & Cognition*, 33, 644–653.

Shaw, P., & McMillion, A. 2008. Proficiency effects and compensation in advanced second-language reading. *Nordic Journal of English Studies*, 7, 123–143.

Simola, J., Holmqvist, K., & Lindgren, M. 2009. Right visual field advantage in parafoveal processing: Evidence from eye-fixation related potentials. *Brain and Language*, 111, 101–113.

Skehan, P. 2003. Task-based instruction. *Language Teaching*, 36, 1–14.

Snellings, P., Van Gelderen, A., & De Glopper, K. 2002. Lexical retrieval: An aspect of fluent second language production that can be enhanced. *Language Learning*, 52, 723–754.

Snellings, P., Van Gelderen, A., & De Glopper, K. 2004. The effect of enhanced lexical retrieval on second language writing: A classroom experiment. *Applied Psycholinguistics*, 25, 175–200.

Stroop, J. R. 1935. Studies of interference in serial verbal reaction. *Journal of Experimental Psychology*, 18, 643–662.

Sumner, E., Connelly, V., & Barnett, A. 2013. Children with dyslexia are slow writers because they pause more often and not because they are slow at handwriting execution. *Reading & Writing: An Interdisciplinary Journal*, 26, 991–1008.

Torppa, M., Georgiou, G., Salmi, P., Eklund, K., & Lyytinen, H. 2012. Examining the double-deficit hypothesis in an orthographically consistent language. *Scientific Studies of Reading*, 16, 287–315.

Vanderwood, M. L., Tung, C. Y., & Checca, C. J. 2014. Predictive validity and accuracy of oral reading fluency for English learners. *Journal of Psychoeducational Assessment*, 32, 249–258.

Van Gelderen, A., Schoonen, R., Stoel, R. D., De Glopper, K., & Hulstijn, J. 2007. Development of adolescent reading comprehension in language 1 and language 2: A longitudinal analysis of constituent components. *Journal of Educational Psychology*, 99, 477–491.

Verhoeven, L. 2000. Components in early second language reading and spelling. *Scientific Studies of Reading*, 4, 313–330.

Walczyk, J. J., Marsiglia, C. S., Johns, A. K., & Bryan, K. S. 2004. Children's compensations for poorly automated reading skills. *Discourse Processes*, 37, 47–66.

Walter, C. 2008. Phonology in second language reading: Not an optional extra. *TESOL Quarterly*, 42, 455–474.

Wolf, M. 1986. Rapid Alternating Stimulus naming in the developmental dyslexias. *Brain and Language*, 27, 360–379.

Wolf, M., & Bowers, P. G. 1999. The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology*, 91, 415–438.

Yamashita, J. 2013. Word recognition subcomponents and passage level reading in a foreign language. *Reading in a Foreign Language*, 25, 52–71.

1 National Board of Education. 2009. National Core Curriculum. Retrieved August 29, 2016, from http://www.oph.fi/download/138886_national_core_curriculum_for_instruction_preparing_for_basic_education_2009.pdf

2 English word list: Leech, G., Rayson, R. & Wilson, A. *Word Frequencies in Written and Spoken English: based on the British National Corpus*. Retrieved October 15, 2010, from http://ucrel.lancs.ac.uk/bncfreq/lists/1_2_all_freq.txt
Russian word list. Retrieved October 15, 2010, from http://corpus.leeds.ac.uk/serge/frqlist/rnc-modern.num.html

3 Word chain text excerpts:
Grade 4
Finnish: Kyllijoki, V-J., Liuskari, H., Rokka, P., Ruuska, H. & Saravesi, S. 2004. *Sanataituri 5*. Otava, p. 135.
English: Westlake, P., Aula, T., Heino, T., Lintunen, M-L., Pitkänen, E-L., Satamo, S. & Seppovaara, A. 1999. *Yes Tales*. Porvoo: WSOY, p. 82–84.
Grade 8
Finnish: Herajärvi, S., Laine, P., Paasio-Leimola, L. & Vartia, P. 2006. *Loitsu Äidinkielen kirjallisuus 9*. "Lapsen Piika" by M. Canth. Otava, p. 156.
English: Westlake, P., Auvinen, M., Kangaspunta, R., Lehtonen, E. & Peuraniemi, J. 2004. *Key English 9*. "Australia". Porvoo: WSOY, p. 97–98.
Grade 11
Finnish: Herajärvi, S., Laine, P., Paasio-Leimola, L. & Vartia, P. 2006. *Loitsu Äidinkielen kirjallisuus 9*. "Lapsen Piika" by M. Canth. Otava, p. 156.
English: Davies, M., Mäkelä, A.-M., Nikkanen, L., Sutela, T., Säteri, L. & Vuorinen, P. 2006. *In Touch, course 8*. "Causing a stir". Porvoo: WSOY, p. 57.

4 Pearson Young Leaners Test: http://pearsonpte.com/pte-young-learners/
Pearson General: http://pearsonpte.com/pte-general/ Retrieved December 21, 2012.

5 National Certificates of Language Proficiency. Retrieved December 21, 2012, from http://oph.fi/english/services/yki

6 National Board of Education. 2004. National Core Curriculum for basic education 2004. Retrieved December 31, 2012, from www.oph.fi/english/curricula_and_qualifications/basic_education

ORIGINAL PAPERS

I

SPEED IN COGNITIVE TASKS AS AN INDICATOR OF SECOND/FOREIGN LANGUAGE READING AND WRITING SKILLS

Olkkonen, S. (2013)

Estonian Papers in Applied Linguistics, 9, 195–208.

doi:10.5128/ERYa9.13

EESTI RAKENDUSLINGVISTIKA ÜHINGU AASTARAAMAT 9, 195–208

SPEED IN COGNITIVE TASKS AS AN INDICATOR OF SECOND/FOREIGN LANGUAGE READING AND WRITING SKILLS

Sanna Olkkonen

Abstract. In a cross-sectional study 823 Finnish school children were tested to examine the relation between speed of performance in cognitive and linguistic tasks and second/foreign language reading and writing. Participants were Finnish-speakers with English as foreign language and Russian-speakers with Finnish as second language which made it possible to compare the results across these two language groups. The Finnish group was furthermore divided into three groups by age to see how speed develops with age and education. Groups were tested with a number of cognitive instruments that included measures of speed of performance. Overall, performance on the speed measures improved with age; often, the second language learners outperformed their foreign language peers of the same age. Regression analyses indicated that speed measures could predict from 20% to over 40% of performance in second/foreign language reading and writing tasks. Prediction was somewhat stronger for writing than reading. The best predictors were also somewhat different for the foreign and second language learners, as well as for the different age groups.*

Keywords: processing speed, reading, writing, cognitive testing, L2, Finnish, English, Russian

1. Introduction

Fluency in language skills can be seen as a combination of speed and accuracy (Grabe 2010, Segalowitz 2010). Speed and accuracy can either be seen as contrasting or complementary skills, and for example dyslexia is considered as a problem with either speed or accuracy. Usually dyslexics encounter problems in one of these skills, and only the most difficult cases have problems in both. In addition to reading problems, dyslexics often have problems with other kinds of automatic constructs, such

* I would like to thank my supervisors Prof. Ari Huhta and Dr. Lea Nieminen for their guidance and feedback. I also would like to thank the two reviewers for their invaluable comments on the manuscript.

as reciting weekdays, months or multiplication tables (Cronin, Carver 1998). Fluent reading requires fast and automatic knowledge of vocabulary. Automatization of language skills is thus an integral part of language learning. When language skills, and especially word retrieval, become automatic, more cognitive capacity becomes available for the higher processes such as holistic text comprehension (Segalowitz 2003). William Grabe (2010: 72) states that: "Effective L1 reading comprehension generally assumes reading fluency – a person reading at a reasonable reading rate, between 250–300 WPM, using very efficient and fast word recognition skills, and combining information from various sources while reading under fairly intense time constraints." Automaticity and its development have been researched widely in first language (L1) environments, but in second language (L2) acquisition the studies are fewer. Also the role of L1 skills in L2 fluency is a question that has evoked discussion (cf. Van Gelderen et al. 2007). The purpose of the current study is to examine the development of automaticity of vocabulary in both L1 (Finnish or Russian) and L2 (English or Finnish), and their relation to second/foreign language reading and writing skills. This examination was conducted in a cross-sectional design using four different age groups with two different language backgrounds: cross-group comparisons were performed on the speed of performance on the similar cognitive and linguistic tasks in both of their languages.

The capacity of working memory is restricted, so the more automatic word retrieval is, the more capacity there is for the use of higher mental processes (Kirby et al. 2003). Overall, the efficient lower-level processes such as memory, attention and word-decoding skills are a prerequisite for the higher ones. Fluent reading and reading comprehension require fast word retrieval. Automatic access to, and retrieval of, word knowledge are required for reading words as wholes and they are essential for being able to get past grapheme-phoneme method of reading. Automatic processes in general are nonconscious, unintentional, involuntary and effortless. They are also more immune to interference from outside sources (Segalowitz 2003, 2010, Grabe 2010). For example, mother-tongue word recognition is involuntary and unavoidable as the so-called Stroop effect shows. In the Stroop test, when the name of a colour is printed in a colour not denoted by the name (e.g., the word 'red' printed in blue ink instead of red ink), naming the colour of the printed word takes longer and is more prone to errors than when the colour of the ink matches the name of the colour (Stroop 1935). As Maryanne Wolf and Patricia Greig Bowers (1999: 418) state, rapid naming can be seen as 'a microcosm of reading'. Working memory is also responsible for the combination and manipulation of new information. If word retrieval is slow, the beginning of the sentence can disappear from memory and integration of information can be hindered. This might mean that there is a certain threshold level in reading speed that should be attained for adequate comprehension (Kirby et al. 2003). Control of attention and retrieval from long-term memory also rely on working memory capacity. Eye-tracking studies have revealed that good readers can control their attention to some extent and concentrate on the relevant parts of the text while skimming (Kaakinen 2004). Effective word retrieval thus has an effect on all levels of reading. In addition to requirements in L1, fluent reading in another language requires direct connections from concepts to L2 vocabulary, and L2 word recognition would be automatic only when direct connections from the print to meanings have been formed. Connections develop slowly with exposure to language, and especially at the early stages of learning to read in L2 the L1 vocabulary also becomes activated (e.g. Kroll, Sunderman 2003; Segalowitz 2003; Hulstijn et al. 2009).

In the *Jyväskylä Longitudinal Study of Dyslexia* (JLD)[1], which started in 1993, the researchers have been following 200 children from birth and have tried to identify the best predictors for diagnosing learning difficulties as early as possible. The results show that phonological awareness and word recognition can, as early as the age of 3.5 years, predict possible reading problems at school age (Puolakanaho et al. 2007). The role of phonological processing as the best predictor for reading proficiency is supported by several other studies (e.g. Torgesen et al. 1997, Bowey et al. 2004). The effect of phonological processing is most notable in languages with very opaque orthography, such as English. In a more transparent language (e.g. Finnish, German or Italian) word recognition has been shown to be a stronger predictor of reading proficiency (Dufva, Voeten 1999, Di Filippo et al. 2005, Landerl, Wimmer 2008). In such languages, word decoding skills develop quickly, and accuracy seems to be close to a maximum level from a very early age. Thus, the ease of word retrieval seems to be an important differentiating factor. Also Grabe (2010: 76) concludes that L2 research, albeit its scarcity, generally supports the importance of fluency in reading comprehension.

The whole process of fluency development in L2 thus needs to be examined more closely. The current study concentrates on the development of fluency in second language (L2) reading and writing. The main interest in the study was to find out whether, in time-pressure tasks, speed as a reflection of a more or less automatic skill could be used to predict L2 reading and writing, which are higher-order skills. This was done by measuring the speed of performance in cognitive and linguistic tasks. These tasks were conducted in both the mother tongue and in the second or foreign language of the participants. This allowed us to compare the possible differences in the second and foreign language learners, as the second language learners live in an authentic language environment and the foreign language is only learned in the school classes. Also the rarer language combinations of Finnish/English and Russian/Finnish can give us new insights on the matter. The cognitive tasks were conducted orally, as Norman Segalowitz (2010), for example, points out that oral fluency is taken to reflect higher rather than lower-level processing, i.e. articulation. The scores from the speed of performance were then correlated with measures of reading and writing in L2 to see if the former could predict performance in the latter. The three different age groups were also compared to see differences in skill development. The design also allowed us to examine whether the speed of performance in L1 had any affect on the L2 skills, which is a much-debated question. The role of L1 for second and foreign language learners is basically different and the differences in the results of this study also reflect this aspect.

The research questions in this study were:

1. Does speed in cognitive tasks develop with age?
2. Can differences in fluency measures be found between second and foreign language learners?
3. Does the speed in time-pressure cognitive and linguistic tasks performed in L1 and L2 predict second/foreign language reading and writing skills?

[1] See: https://www.jyu.fi/ytk/laitokset/psykologia/huippututkimus/en/research/JLD_main/JLD (21.12.2012).

2. Method

2.1. Participants

The data were gathered in the DIALUKI project (Diagnosing reading and writing in a second or foreign language[2]) in a cross-sectional setting in 2010–2011. Participants (*n* = 823) were selected from 111 schools in Finland. The first group consisted of Finnish-speaking learners of English and was divided into sub-groups from three different grades. All these children had started learning English as a foreign language in the third grade. The second group consisted of primary school children with Russian as L1 and Finnish as L2 (see Table 1). Russian students were immigrants living in Finland, and their language skills varied widely in both languages. At the time of the study, they either participated in preparatory classes or were already integrated within mainstream education, depending on their language skills and length of residence. For some of them Finnish was even a stronger language than Russian. Agreements for participation in the study were collected from the county, schools, parents and children themselves.

Table 1. Participants of the study

Mother tongue	Grade / age	Foreign / second language	
		English	Finnish
Finnish	4th grade / 9–10 years 8th grade / 13–14 years Gymnasium / 17–18 years	210 208 219	–
Russian	3rd–6th grades (mean age 10.9 years)	–	186
Total		637	186

2.2. Procedure

Tests were administered in the school during school hours. There were two parts in the test battery: reading and writing tasks and the speed-measuring linguistic tasks were performed in groups in the classroom; the individual cognitive tasks were computer-assisted and administered by trained assistants.

2.3. Speed-measuring linguistic tasks

Word segmentation task. The children were given a short text written without spaces between the words and without punctuation. The task was to segment the text by drawing a pencil stroke on word boundaries (i.e. between the words). The task was performed in both L1 and L2, but the texts were different for the different age groups. The time the children took to complete the task was recorded (to the nearest 30 seconds); also, the number of correctly separated words was counted.

[2] See: https://www.jyu.fi/hum/laitokset/solki/tutkimus/projektit/dialuki (21.12.2012).

Spelling error detection task. In the spelling error detection task, the children were given a paper with a list of 100 Finnish words each containing a spelling error. The spelling errors were of several different types; either there was a letter missing or extra in the word (*lauvantai* instead of *lauantai*), or there was a wrong letter (*heunäkuu – heinäkuu*), or incorrect inflections (*aurinkonkukka – auringonkukka*). Children were instructed to mark all the spelling errors they could detect but not to correct them. They were given 3.5 minutes to mark as many as they could. The maximum score was 100 points and the number of correct answers was counted. The task was designed for youths and adults, so only the eighth graders and gymnasium students took this task (Holopainen et al. 2004). The task was conducted only in Finnish.

2.4. Cognitive tasks

Students performed the cognitive tasks with the *Cognitive Workshop* computer program developed in the *Jyväskylä Longitudinal Study of Dyslexia*. The test battery consisted of eight different tasks in total, and three of these measured performance speed. These tasks were *Rapidly Presented Words*, *Reading Aloud Word List*, and *Rapid Alternating Stimulus*. The original tasks had been developed and validated in the JLD project; the Russian and English versions of these tasks were developed by the DIALUKI project. Each task is described in more detail below.

Rapidly Presented Words (RPW). The task measured reading (and perceiving) a word as a whole. The words appeared one at a time on the computer screen for 80 milliseconds and the child had to read them aloud. There were 14 words in L1 and 8 or 12 words in L2, depending on students' age. All the words were simple, everyday words but in the task they became progressively longer, from two-letter words up to a nine-letter word. A star first flashed on the screen as a prompt, after which the target word appeared. After 80 milliseconds it was replaced by a mask (e.g., B/W#Q?) to erase the visual image from the working memory. The next stimulus was given only after the child had answered or confirmed not having perceived the word. Correct answers were counted and converted to percentages to enable the comparisons of performances on tasks consisting of a different number of stimuli reported later in this article.

Reading Aloud Word List. The children were given a list consisting of 105 words. They were instructed to read the list as fast and as accurately as they could in one minute. The test was conducted in both L1 and L2. The Finnish word list was from the Lukilasse test battery developed for diagnosing reading in L1 Finnish in primary school (Häyrinen et al. 1999). The English and Russian word lists were designed by sampling from frequency lists for these two languages.[3] Time was measured with a stopwatch and the last word read by the child was marked. Afterwards, the words were divided into syllables, which makes the word lists across languages more (although not completely) comparable. This was necessary as the words in both Finnish and Russian typically consist of several syllables and, thus, take longer to pronounce, whereas many frequent English words are rather short.

Rapid Alternating Stimulus (RAS) and *Rapid Automatic Naming* (RAN) are considered to be very good tests for measuring the automaticity of language skills

[3] See: http://ucrel.lancs.ac.uk/bncfreq/lists/1_2_all_freq.txt (15.10.2010), http://corpus.leeds.ac.uk/serge/frqlist/rnc-modern.num.html (15.10.2010).

(e.g. Cronin, Carver 1998). It is a matrix consisting of letters, numbers, colours or simple objects or a combination of these. The matrix used in this study was a mixed one with numbers, colours and letters or pictures. This made the test more difficult and therefore more prone to mistakes compared with a task consisting of only numbers, for example. The basic matrix consisted of 50 units. For primary school participants it was shorter in L2 (30 units) and letters were replaced by pictures of easily identifiable objects (i.e., pen, car and house). Children were told to read the entire matrix aloud as fast and as accurately as they could. The completion time for the test was measured by a stopwatch. To make comparisons across the 30 and 50-unit versions of the task possible, the overall completion times were converted to time-per-unit values.

2.5. Reading and writing measures

The second/foreign language reading and writing tasks that were in the study were standardised and age-appropriate.

Reading tasks. In the primary school, reading comprehension tasks in English as a foreign language were three tasks from the Pearson Young Learners test. The older students had four tasks from the Pearson Test of English General test. The Finnish L2 reading comprehension tasks were three tasks from the YKI test (National Certificate for Language Proficiency) and ten tasks from DIALANG. All reading tasks were paper-and-pencil tests, except DIALANG, which is an online diagnostic system, designed to assess proficiency in fourteen different European languages. The reading test scores were analysed with Rasch software (Winsteps)[4], and reading ability values were used as dependent variables in the analyses focusing on reading.

Writing tasks. The writing tasks were paper-and-pencil tasks. The English L2 writing tasks were a letter or e-mail to a friend, a short story and an opinion; there was one task for the fourth graders and three for the older groups. Finnish L2 tasks included a letter to a friend and a short story. The tasks were taken from the above-mentioned Pearson tests and the Topling research project. Two professional raters evaluated the learners' performances against the Common European Framework of Reference (CEFR) 6-point scale[5]. The ratings were analysed with the multi-faceted Rasch programme Facets[4], and the ability values that take into account both the raters' leniency/severity and the task difficulty were later used as the dependent variables in the regression analyses reported below.

3. Results

3.1. Analysis of variance

First the speed of performance in the cognitive tasks was compared across the age groups (fourth graders, eighth graders, and gymnasium students). The second and foreign language learners (i.e. the Russian-speaking learners of Finnish and Finnish-speaking learners of English) were also compared to see the possible differences

[4] See: http://www.winsteps.com (21.12.2012).
[5] See (21.12.2012): YKI test http://www.oph.fi/english/mobility/testing_language_skills; DIALANG http://www.lancs.ac.uk/researchenterprise/dialang/about; Topling. Paths in Second Language Acquisition https://www.jyu.fi/hum/laitokset/kielet/topling/en.

between these groups. For these, one-way analysis of variance (ANOVA, see Hatch, Farhady 1982) was used.

The results concerning the *Rapidly Presented Words* (RPW) task are presented for L1 and L2, respectively. Tables 1 and 2 present the means (M) and standard deviations (SD) of the percentages of correct words in each group. The results show that automatic recognition of vocabulary increased with age. The RPW task was able to distinguish between different age groups: a particularly clear distinction was found between the primary school and the older children both in the L1 and L2 condition (for L1 Finnish/Russian $F = 75.38$, $p < 0.0001$, $df = 3$, and for L2 English/Finnish $F = 145.03$, $p < .0001$, $df = 3$. The effect size η^2 for L1 was 0.235 and for L2 0.364). The task was quite easy for the two oldest age groups and a ceiling effect was in fact found for them; thus, the RPW tasks used in the study could not differentiate between the eighth graders and the gymnasium students. Interestingly, in the L1 task, both the Finnish and Russian-speaking children performed equally well (assuming that the different language versions of the task are comparable; see Table 2), whereas in the L2 task, the Russian-speakers outperformed their same age Finnish-speaking peers although they fell far behind the older Finnish-speaking learners (Table 3).

Table 2. Rapidly Presented Words in L1 (Finnish / Russian): percentage of correct words out of 14

Group	n	M	SD
4th grade	203	73.65	24.91
Russian primary	153	70.63	31.88
8th grade	194	93.37	6.32
Gymnasium	191	95.10	4.89

Table 3. Rapidly Presented Words in L2 (English / Finnish): percentage of correct words out of eight words for primary school groups and twelve words for older groups

Group	n	M	SD
4th grade	202	58.79	27.83
Russian primary	177	66.24	32.33
8th grade	192	93.45	11.06
Gymnasium	192	97.35	8.56

The results from the *Reading Aloud Word List* task are displayed as syllables read per second (see Tables 4 and 5). The results show that speed (syllables per second) increased with age and, presumably, with language skill development. Overall, the results are somewhat similar to those reported above for the *Rapidly Presented Words* task. On the whole, the groups are very clearly separated (for L1 Finnish/Russian $F = 230.66$, $p < 0.0001$, $df = 3$; and for L2 English/Finnish $F = 210.08$, $p < 0.0001$, $df = 3$. The effect sizes are very large: 0.474 for L1 and 0.452 for L2). Again, the primary school age students performed less well than the older students but this time the Russian-speaking students' L2 Finnish performance was not only better than their same-age Finnish peers' L2 English performance but was at the same level as the eighth graders (Table 5). On this task, the oldest group (gymnasium) outperformed all younger groups statistically significantly, both in the L1 and L2 tasks.

Table 4. Reading Aloud Word List in L1 (Finnish / Russian): syllables per second

Group	n	M	SD
4th grade	205	2.76	0.76
Russian primary	182	2.60	1.20
8th grade	195	4.13	0.97
Gymnasium	189	4.86	1.01

Table 5. Reading Aloud Word List in L2 (English / Finnish): syllables per second

Group	n	M	SD
4th grade	204	1.32	0.42
Russian primary	182	2.30	0.85
8th grade	194	2.36	0.66
Gymnasium	190	2.95	0.66

In the *Rapid Alternating Stimulus* (RAS) task, the performance speed also increased with age (see Tables 6 and 7). The results of the RAS tasks were almost identical to those reported above for the *Reading Aloud Word List*: the two younger groups (Finnish-speaking and Russian-speaking primary students) performed the same on the L1 RAS task but the Russian-speakers outperformed the Finns on the second/ foreign language version of the same task, and reached the same average speed as the eighth-grade Finns. Again, the gymnasium students did these tasks significantly faster than any of the younger groups. (Details of the ANOVA results are as follows: for L1 $F = 82.99$, $p < 0.0001$, $df = 3$; and for L2 $F = 120.55$, $p < 0.0001$, $df = 3$. The effect sizes are also quite large: 0.246 for L1 and 0.321 for L2).

Table 6. Rapid Alternating Stimulus in L1 (Finnish/Russian): average time in seconds spent on each item (50 items)

Group	n	M	SD
4th grade	204	0.87	0.18
Russian primary	177	0.83	0.31
8th grade	195	0.68	0.16
Gymnasium	190	0.59	0.13

Table 7. Rapid Alternating Stimulus in L2 (English/Finnish): average time in seconds spent on each item (30 items in primary school and 50 items in 8th grade and gymnasium)

Group	n	M	SD
4th grade	205	1.20	0.41
Russian primary	180	0.89	0.22
8th grade	195	0.86	0.25
Gymnasium	190	0.68	0.15

3.2. Regression analyses

Stepwise multiple regression analyses were used to test to what extent the time-pressure cognitive and linguistic tasks explain reading and writing skills. The dependent variables were reading and writing in L2 (ability measures derived from the Rasch analyses of the reading scores and ratings of the writing scripts) and the independent variables (IV) were the time-pressure linguistic and cognitive tasks (see Table 8).

Overall, the speed-related measures predicted writing in L2 better than reading in L2: The amount of explained variance was at least 10%, sometimes even 20%, higher for writing. For reading, about 20% of the variance could be predicted from performance on the speed measures included in the study. For writing, the explained variance rose to about 30–40%. Interestingly, the prediction was somewhat better for the L2 (Russian-speakers) learners than it was for the foreign language learners (Finnish-speakers), in both reading and writing.

For the Finnish-speaking learners of English, the best predictor for the primary school group was shown to be L2 the *Rapidly Presented Words* (RPW) either in L1 or L2, and *Rapid Alternating Stimulus* (RAS) in L2 for other grades. However, several other measures of speed, including some linguistic speed tasks, also predicted reading and writing performance significantly among the Finnish-speaking learners. In comparison, the set of predictors of Finnish as a L2 reading and writing among the Russian-speaking learners looks somewhat different. *Reading Aloud Word List* in L2 Finnish dominates as the best predictor for both reading and writing. *Rapidly Presented Words* in L2 and (for writing) also *Reading Aloud Word* List in L1 Russian appear as significant predictors of L2 macro skills.

Table 8. Speed in the time-pressure cognitive tasks as a predictor of reading and writing in a second/foreign language

Age groups	Dependent variable	Adjusted R Squared	% variance	1st IV	2nd IV	3rd IV	4th IV
4th grade Finnish L1	Writing L2	0.272	**27.2%**	RPW L1	RAS L2		
	Reading L2	0.188	**18.8%**	RPW L2	RAS L2		
Primary Russian L1	Writing L2	0.459	**45.9%**	Word List L2	Word List L1	RPW L2	
	Reading L2	0.300	**30.0%**	Word List L2	RPW L2		
8th grade Finnish L1	Writing L2	0.422	**42.2%**	RAS L2	Segmen-tation L1	RPW L2	Segmen-tation L2
	Reading L2	0.218	**21.8%**	RAS L2	RPW L2		
Gymnasium Finnish L1	Writing L2	0.322	**32.2%**	RAS L2	Spelling errors L1	Word List L1	
	Reading L2	0.242	**24.2%**	RAS L2	Segmen-tation L1	Word List L1	

4. Discussion

The study shows that speed in the fluency tasks increases with age and education, at least among the learner groups studied here. Three different age groups were distinguished in most tasks. In both primary school groups, in which second or foreign language skills are only starting to develop, performance on the L1 versions of the tasks was always better than in the L2 tasks. When language skills become more fluent, the difference starts to diminish. In this study the difference was still mostly visible also in the older groups, which is in accord with Jan Hulstijn, Amos Van Gelderen and Rob Schoonen (2009), for example. Where no difference between L1 and L2 could be found, in particular in the *Rapidly Presented Words* task, this was mainly due to the task construction. The English version of the RPW task was clearly too easy for the older students and the results were influenced by a ceiling effect. In the *Word List* task the Finnish words used were much longer with complex inflections while the English and Russian versions contained easier basic words. For the comparability of the results they were analysed as syllable/second. Interestingly, the Finnish version of the list was also the language version that did have the most predictive value with reading and writing skills in the regression analysis, as both L1 and L2 task. It was quite a challenging task even for older children as the word-length came up to a 22 letter-word (e.g. *prosessikirjoittaminen* 'process writing').

The analyses revealed some differences between the two language groups. The primary school Russian speakers were closest to the corresponding Finnish age group in L1 tasks, but in L2 tasks they were markedly better and even indistinguishable from the eighth grader Finns. This might be caused by several factors. For some learners labelled as Russian-speakers, Finnish was in fact their stronger language. Their Russian skills were more varied, as we can see from the standard deviations on their L1 tasks. For the Finnish students English is a foreign language learned at school; the youngest group had only studied English for slightly over a year. For the Russian speakers, Finnish is a second language learned in an authentic environment. There is also a slight age difference in favour of Russian speakers. Positive effects that are often associated with bilingualism can also have some effect on the results (e.g. Bialystok 2001).

The time-pressure tasks studied here were found to predict the learners' reading and writing performance in L2 rather well. The regression analyses indicated that the speed measures predicted writing in L2 better than reading in L2, and that the prediction was somewhat better for the second language group than it was for the foreign language groups. The best predictor of reading and writing measures among the Finnish-speaking primary school group was the *Rapidly Presented Words*, which was also a significant predictor for the Russian-speaking primary school learners. The current versions of the RPW task, which made use of simple, very frequent words were too easy for the older students whose word recognition skills could be expected to be more automatic. It is possible that a more difficult version of the RPW could be useful in predicting L2 macro skills such as reading and writing but this remains to be studied in the future. Also, in line with earlier studies, *Rapid Alternating Stimulus* explained a significant amount of variance in L2 reading and especially writing, at least for the Finnish speakers. It can be argued to be a good predictor of automaticity in language as it combines different skills that are required for fluent reading. Overall, the results are in accordance with previous research in that the lower level processes can predict more holistic processes, such as reading and writing also in a foreign language. In addition to the L2 tasks, the L1 tasks also had predictive power in the current study which has not been the case in all studies (cf. Van Gelderen et al. 2007). It would be reasonable to assume that L1 had influence only on the lower grades, when the L2 skills are still at quite a low level. This study shows the L1 effect on the L2 skills also in the older groups. For them it was especially the L1 linguistic tasks (i.e. *Segmentation* and *Spelling errors*) that showed up in the regression analysis.

Hulstijn et al. (2009) conclude their analysis stating that processing speed may not be a very useful criteria for automaticity. Segalowitz (2010: 79) also proposes that for a skill to be regarded as automatized, speed cannot be the only criterion. This study has looked only at the speed component of fluency, and to examine the wider phenomenon of fluency, an analysis of accuracy of performance could bring a more balanced and informative view. From the data of this study it was evident that often the time to complete the task was only a part of the whole performance. The time-pressure in the tasks seemed to result in a trade-off of speed and accuracy, with quick performance sometimes revealing on a closer inspection to consist only of approximated responses. The combined effect of speed and accuracy should be therefore examined in the further analyses.

5. Conclusion

The main conclusion from the work presented here is that the speed in the time-pressure tasks in both L1 and L2 can be used in predicting L2 reading and especially writing skills. This applied also to the older students, even when some of the tasks used in this study were not very suitable for them. All in all, the tasks with the best predictive value differed somewhat for the different age groups and for the second or foreign language learners. For further study, an even more informative view could be attained by combining these speed of performance measures with accuracy measures from the same tasks.

Abbreviations

L1 first language
L2 second or foreign language
M mean
RAS *Rapid Alternating Stimulus* test
RPW *Rapidly Presented Words* test
SD standard deviation

References

Bialystok, Ellen 2001. Bilingualism in Development. Language, Literacy and Cognition. Cambridge: Cambridge University Press.

Bowey, Judith A.; Storey, Tiffany; Ferguson, Angela N. 2004. The association between continuous naming speed and word reading skill in fourth- to sixth-grade children. –

Australian Journal of Psychology, 56 (3), 155–163. http://dx.doi.org/10.1080/00049530412331283345
Cronin, Virginia; Carver, Paula 1998. Phonological sensitivity, rapid naming, and beginning reading. – Applied Psycholinguistics, 19 (3), 447–461. http://dx.doi.org/10.1017/S0142716400010262
Di Filippo, Gloria; Brizzolara, Daniela; Chilosi, Anna; De Luca, Maria; Judica, Anna; Pecini, Chiara; Spinelli, Donatella; Zoccolotti, Pierluigi 2005. Rapid naming, not cancellation speed or articulation rate, predicts reading in an orthographically regular language (Italian). – Developmental Neuropsychology, 30, 885–904.
Dufva, Mia; Voeten, Marinus J. M. 1999. Native language literacy and phonological memory as prerequisites for learning English as a foreign language. – Applied Psycholinguistics, 20 (3), 329–348. http://dx.doi.org/10.1017/S014271649900301X
Grabe, William 2010. Fluency in reading – thirty-five years later. – Reading in a Foreign Language, 22, 71–83.
Hatch, Evelyn; Farhady, Hossein 1982. Research Design and Statistics for Applied Linguistics. Cambridge: Newbury House.
Holopainen, Leena; Kairaluoma, Leila; Nevala, Jukka; Ahonen, Timo; Aro, Mikko (Eds.) 2004. Lukivaikeuksien seulontamenetelmä nuorille ja aikuisille. [Dyslexia Screening Test for Youth and Adults.] Jyväskylä: Niilo Mäki Instituutti.
Hulstijn, Jan; Van Gelderen, Amos; Schoonen, Rob 2009. Automatization in second-language acquisition: What does the coefficient of variation tell us? – Applied Psycholinguistics, 30 (4), 555–582. http://dx.doi.org/10.1017/S0142716409990014
Häyrinen, Taru; Serenius-Sirve, Silve; Korkman, Marit 1999. Lukilasse. Lukemisen, kirjoittamisen ja laskemisen seulontatestistö peruskoulun ala-asteen luokille 1–6. [The Lukilasse Graded Achievement Package for Comprehensive School Age Children.] Helsinki: Psykologien Kustannus Oy.
Kaakinen, Johanna K. 2004. Perspective effects on text comprehension: Evidence from recall, eyetracking, and think-alouds. Annales Universitatis Turkuensis, Ser. B, 275.
Kirby, John R.; Parrila, Rauno K.; Pfeiffer, Shannon L. 2003. Naming speed and phonological awareness as predictors of reading development. – Journal of Educational Psychology, 95 (3), 453–464. http://dx.doi.org/10.1037/0022-0663.95.3.453
Kroll, Judith F.; Sunderman, Gretchen 2003. Cognitive processes in L2 learners and bilinguals. – Catherine J. Doughty, Michael H. Long (Eds.). Handbook of Second Language Acquisition. Malden, MA: Blackwell Publishing.
Landerl, Karin; Wimmer, Heinz 2008. Development of word reading fluency and orthographic spelling in a consistent orthography: An 8-year follow-up. – Journal of Educational Psychology, 100 (1), 150–161. http://dx.doi.org/10.1037/0022-0663.100.1.150
Puolakanaho, Anne; Ahonen, Timo; Aro, Mikko; Eklund, Kenneth; Leppänen, Paavo; Poikkeus, Anna-Maija; Tolvanen, Asko; Torppa, Minna; Lyytinen, Heikki 2007. Very early phonological and language skills: estimating individual risk of reading disability. – Journal of Child Psychology and Psychiatry, 48 (9), 923–931. http://dx.doi.org/10.1111/j.1469-7610.2007.01763.x
Torgesen, Joseph K.; Wagner, Richard K.; Rashotte, Carol A.; Burgess, Stephen; Hecht, Stephen 1997. Contributions of phonological awareness and rapid automatic naming ability to the growth of word-reading skills in second- to fifth-grade children. – Scientific Studies of Reading, 1 (2), 161–185. http://dx.doi.org/10.1207/s1532799xssr0102_4
Segalowitz, Norman 2003. Automaticity and second languages. – Catherine J. Doughty, Michael H. Long (Eds.). Handbook of Second Language Acquisition. Malden, MA: Blackwell Publishing, 382–408.
Segalowitz, Norman 2010. The Cognitive Bases of Second Language Fluency. New York: Routledge.

Stroop, John Ridley 1935. Studies of interference in serial verbal reaction. – Journal of Experimental Psychology, 18 (6), 643–662. http://dx.doi.org/10.1037/h0054651
Van Gelderen, Amos; Schoonen, Rob; Stoel, R. D.; De Glopper, Kees; Hulstijn, Jan 2007. Development of adolescent reading comprehension in language 1 and language 2: A longitudinal analysis of constituent components. – Journal of Educational Psychology, 99 (3), 477–491. http://dx.doi.org/10.1037/0022-0663.99.3.477
Wolf, Maryanne; Bowers, Patricia Greig 1999. The double-deficit hypothesis for the developmental dyslexias. – Journal of Educational Psychology, 91 (3), 415–438. http://dx.doi.org/10.1037/0022-0663.91.3.415

Online references

DIALANG. http://www.lancs.ac.uk/researchenterprise/dialang/about (21.12.2012).
DIALUKI = Diagnosing reading and writing in a second or foreign language. https://www.jyu.fi/hum/laitokset/solki/tutkimus/projektit/dialuki (21.12.2012).
English frequency lists. http://ucrel.lancs.ac.uk/bncfreq/lists/1_2_all_freq.txt (15.10.2010).
JLD = Jyväskylä Longitudinal Study of Dyslexia. https://www.jyu.fi/ytk/laitokset/psykologia/huippututkimus/en/research/JLD_main/JLD (21.12.2012).
Russian frequency lists. http://corpus.leeds.ac.uk/serge/frqlist/rnc-modern.num.html (15.10.2010).
Topling. Paths in Second Language Acquisition. https://www.jyu.fi/hum/laitokset/kielet/topling/en (21.12.2012).
YKI test = National Certificate of Language Proficiency. http://www.oph.fi/english/mobility/testing_language_skills (21.12.2012).
Winsteps program. http://www.winsteps.com (21.12.2012).

Sanna Olkkonen (University of Jyväskylä). Her academic research interests include cognitive processes in second language acquisition, especially development of fluency in second language.
sanna.olkkonen@jyu.fi

KOGNITIIVSETE ÜLESANNETE LAHENDAMISE KIIRUS KUI VÕÕRKEELE / TEISE KEELE LUGEMIS- JA KIRJUTAMISOSKUSE NÄITAJA

Sanna Olkkonen
Jyväskylä Ülikool

Artikkel analüüsib võõrkeeleoskuse automaatsuse arengut, mis on teise keele uuringutes veel vähekäsitletud valdkond. Käesolev läbilõikeuuring kasutas ajasurvega kognitiivseid ja keelelisi ülesandeid ning analüüsis, kuidas soorituskiirus vanusega ning õppimise käigus areneb. Andmed koguti aastatel 2010–2011 Soome koolide 823 õpilaselt. Soomekeelsed lapsed, kes õppisid põhikooli neljandas ja kaheksandas klassis ning gümnaasiumi teises klassis, õppisid inglise keelt võõrkeelena. Venekeelsed lapsed olid algkooliõpilased ning õppisid soome keelt teise keelena. Hindamaks kiirust mõõtvate ülesannete võimet ennustada võõrkeele (teise keele) lugemis- ja kirjutamisoskust, kasutati regressioonanalüüsi.

Uuring näitas, et algkoolis, kui keeleoskus alles hakkab arenema, oli sooritus emakeeles parem kui teises keeles või võõrkeeles. Kui keeleoskus enam automatiseerus, hakkas see erinevus kaduma. Soorituskiirus üldiselt paranes vanusega, kuigi kohati ei olnud võimalik kahe vanema soomekeelse rühma tulemusi eristada. Venekeelsete algkooliõpilaste tulemused olid samas vanuses soomekeelse rühma omadest sageli paremad. Siiski ei olnud need enamasti vanemate soomekeelsete õpilaste tasemel. Parimateks võõrkeele lugemis- ja kirjutamisoskuste ennustajateks osutusid nooremate õpilaste puhul sõnade kiire esitamise ülesanne (*Rapidly Presented Words*) ja vanemate õpilaste puhul kiirelt vahelduv stiimul (*Rapid Alternating Stimulus*). Venekeelsete õpilaste puhul oli parimaks ennustajaks sõnaloendi valjult ettelugemise ülesanne (*Reading Aloud Word List*). Ka emakeelsed ülesanded suutsid teatud määral võõrkeele (teise keele) oskust ennustada. Kokkuvõttes ennustasid ajasurve ülesanded paremini kirjutamist kui lugemist.

Võtmesõnad: soorituskiirus, lugemine, kirjutamine, kognitiivne testimine, teine keel, soome keel, inglise keel, vene keel

II

FLUENCY OF LEXICAL ACCESS IN L1 AND L2 IN RELATION TO L2 LITERACY

Olkkonen, S., Eklund, K. & Leppänen, P. H. T.

Manuscript

III

PROCESSING LIMITATIONS IN L2 FLUENCY: ANALYSIS OF INACCURACIES IN LEXICAL ACCESS

Olkkonen, S. (2017)

APPLES: Journal of Applied Language Studies, 11, 19–41.

Apples – Journal of Applied Language Studies
Vol. 11, 1, 2017, 19–41

Processing limitations in L2 fluency: Analysis of inaccuracies in lexical access

Sanna Olkkonen, University of Jyväskylä

Oral fluency is widely included in second language assessments, but its relationship to language proficiency is not straightforward. In the current study, data gathered in an experimental setting were examined with an exploratory fluency analysis. The aim was to examine the relationship between fluency of lexical access and proficiency in foreign language (L2). Fluency of the lexical access was studied by analysing inaccuracies in one word recognition and one word retrieval task. To see if proficiency had an effect on the number and the type of inaccuracies, lexical access tasks were carried out for 563 Finnish school children from grades 4, 8, and 11 in their L2 (English). Proficiency in L2 was expected to develop during school education. The inaccuracies were proposed to stem from processing limitations in language use, i.e., inefficiency of lexical access, or from control of attention. The hypothesis was that if lexical access is not automatized, there are less resources for attention-control in recognising and retrieving words. Therefore, the inaccuracies in L2 relating to inefficiency were hypothesised to decrease with proficiency, whereas the ones relating to control of attention were proposed to be more stable or to increase. Furthermore, the fluency of L1 lexical access was used as a control measure. The results offered some confirmation to these hypotheses. For example, some evidence for more available resources in correcting and monitoring speech was found for the older students. The overall results highlight caution in assessing L2 fluency, as not all types of inaccuracies were connected with lower proficiency.

Keywords: fluency, oral proficiency, information-processing, lexical access

1 Introduction

Second language (L2) fluency has received growing interest in recent years (e.g., Akamatsu, 2008; Bosker, Pinget, Quené, Sanders & De Jong, 2013; De Jong, Groenhout, Schoonen & Hulstijn, 2015; Engelhardt, Corley, Nigg & Ferreira, 2010; Housen & Kuiken, 2009; Kahng, 2014; Kame'enui & Simmons, 2001; Riggenbach, 2000; Segalowitz, 2010; Snellings, Van Gelderen & De Glopper, 2002). No clear consensus has been reached, though, on what is meant by fluency of language use. Despite this, oral fluency is widely used in assessing L2 proficiency; it is, for example, one of the five assessment criteria in the Common European Framework of Reference for Languages (CEFR, Council of Europe, 2007; see also Koponen & Riggenbach, 2000). The criteria given for assessing fluency are very wide

Corresponding author's email: sanna.olkkonen@jyu.fi
ISSN: 1457-9863
Publisher: Centre for Applied Language Studies
University of Jyväskylä

http://apples.jyu.fi
http://dx.doi.org/10.17011/apples/urn.201704252081

VERTAISARVIOITU
KOLLEGIALT GRANSKAD
PEER-REVIEWED
www.tsv.fi/tunnus

generalisations, and CEFR defines spoken fluency in terms of pausing and 'natural, smooth flow of language'. This is problematic, as it is questionable whether even L1 speech can be considered as 'naturally smooth' when it includes pausing, self-corrections, and repetitions as well (Lennon, 1990). Furthermore, not all of the aspects of oral fluency seem to be as indicative of proficiency level as others. As De Jong and colleagues (2015) showed, the duration of pauses is more closely connected to L1 speaking style than to L2 proficiency. Therefore, duration of pauses should be considered only to a modest extent when assessing L2 fluency. This highlights the importance of understanding, which are the features of oral fluency that are most reliably indicative of L2 proficiency.

In the current study, it is proposed that oral fluency and its relationship to L2 proficiency might be studied from the viewpoint of processing limitations. For the beginning language learners, even the most basic aspects of language, such as phonological coding and lexical search require much of the resources. When these processes start to automatize with proficiency, the resources can be directed to higher aspects of language, such as more global planning and comprehension monitoring. Fluent language use thus means a balance between the *automatic* and *controlled* processes (Schmidt, 1992; Segalowitz, 2000, 2010, p. 91). Examining these two levels of processing in L2 use may reveal us important information on how fluency and proficiency are connected. Therefore, it is proposed that the *inaccuracies* in language production may offer a fruitful way of tapping into these processes. Fluency was measured with two standardised lexical access tasks (measuring word recognition and word retrieval), and in addition to L2, these tasks were conducted also in L1 to use its fluency as a control variable. To be able to examine the differences between proficiency levels, the tasks were conducted for 580 Finnish schoolchildren from three different L2 proficiency levels: Grade 4 (in primary school, age 10), Grade 8 (in lower secondary school, age 14), and Grade 11 (in upper secondary school, age 17). The use of grade levels as a proxy for the developing proficiency has been previously confirmed with L2 writing tasks, assessed along the CEFR 6-point scale (see Olkkonen, Eklund & Leppänen, submitted).

1.1 Defining and measuring fluency

The range of definitions of L2 fluency is wide. It can mean the overall proficiency of a language; only the speed of language production; the ease or efficiency of production; or including prosodically appropriate expression and phrasing (for overviews: Housen & Kuiken, 2009; Koponen & Riggenbach, 2000). The most comprehensive model to date, however, is Segalowitz's (2010) distinction between cognitive, utterance, and perceived fluency. The model builds on Lennon's division (1990) of speech production and underlying psycholinguistic processes, where the surface fluency is a reflection of the underlying cognitive fluency. *Perceived fluency* refers to the interpretations the listeners make on a speaker's proficiency based on their surface fluency (cf. Bosker et al., 2013). *Utterance fluency* refers to the surface structure that reflects the cognitive level and is directly measurable. Utterance fluency is often operationalised as the rate of repairs, pauses, and speed (Skehan, 2003). Pausing and speed have been shown to be relatively good markers of (dis)fluency (Bosker et al., 2013; De Jong et al., 2015; Freed, 2000; Kahng, 2014). On the other hand, the repairs (false starts, repetitions, replacements, and reformulations) have been studied to a much lesser extent, and their relationship to L2 proficiency has been questioned (Engelhardt et al., 2010;

Gilabert, 2007; Kormos, 1999; Lennon, 1990). These results have important implications for the current study as well, when considering the different types of disfluencies (inaccuracies) and proficiency.

Cognitive fluency refers to how efficiently (i.e., quickly and accurately) the processes underlying language production operate (Segalowitz, 2010, p. 48). These underlying processes include decoding words and lexical access (Grabe, 2009). Our cognitive resources are limited, and thus, the more fluently these underlying processes operate, the more smoothly the higher-level processes can be executed (Kirby, Parrila & Pfeiffer, 2003; LaBerge & Samuels, 1974). These higher-level processes include, e.g., monitoring and comprehension. In Segalowitz's formulation, cognitive fluency consists of two complementary processes: efficiency and attention-control (2000, 2010, p. 91). Efficient language use (whether reading, speaking, writing or listening) entails fast and accurate processing. When this processing is inefficient, there are fewer resources available for the attention-related processes (Kahng, 2014; Kormos, 2000; Walczyk, Marsiglia, Johns & Bryan, 2004). This means, for example, that a speaker is not able to efficiently suppress irrelevant information, shift between different requirements of the situation, or monitor one's performance (see Segalowitz, 2010, p. 93). These skills are very important for language use, as Segalowitz and Frenkiel-Fishman (2005) showed that the efficiency of L1 and L2 task-switching accounted for 59% of the variance in L2 proficiency.

Studies aimed at cognitive fluency measurements are rare, and often highly indirect measures are used, such as stimulated recall settings (Kahng, 2014). One possibility for a more direct measurement has been proposed to be the efficiency of access to lexicon (Segalowitz, 2010, p. 75; Segalowitz & Frenkiel-Fishman, 2005). Even though often measured orally, the efficiency of lexical access has been shown to be a good reflection of fluency of mental processing (Denckla & Cutting, 1999; Di Filippo et al., 2005). Compared to longer monologue or dialogue settings, the reduced contextual and communicative aspects of the lexical access measurements may help to tap into the cognitive processes in a more ecologically valid way. The term *lexical access* has been used as both referring to word retrieval (e.g., Gholamain & Geva, 1999; Kirby et al., 2003), and to word recognition (e.g., Akamatsu, 2008; Grabe, 2009). On the other hand, they might be both interpreted as related processes, both tapping into lexical access but from different directions (Snellings et al., 2002). Both word retrieval and recognition rely on retrieving a phonological form of words, while utilising different subcomponents of lexical access (see also Gollan et al., 2011). In the current study, this approach was applied by measuring fluency in both word recognition and retrieval.

1.2 Inaccuracies and processing limitations in lexical access

Lennon, in his account on L2 fluency, stated that "for most speakers in most situations, processing demands, rather than deficient knowledge, will limit fluency" (2000, p. 27). Therefore, studying the connections between L2 fluency and proficiency might be performed by tapping the processing demands. Especially, the processing limitations in language use may prove a fruitful source of information. As Just and Carpenter's (1992) influential theory of information processing defines, human cognition is limited and these limits are set by the working memory capacity. When our language skills are in development, the cognitive resources are directed by necessity to low-level processes, such as lexical access. There is not enough capacity for the higher processes, like comprehension or correcting your speech, which may result as errors in language production. These errors, therefore, serve as reflections of the proficiency level. On the other hand, Ehri (1991) has hypothesised that certain types of errors could be connected with the development of language skills, so that they appear only after certain proficiency threshold. Therefore, the breakdowns of fluency may also be indications of developing proficiency, and, as such, they can offer important insights into the influence of processing capacity in the language use (see e.g., Engelhardt et al., 2010; Protopapas, Fakou, Drakopoulou, Skaloumbakas & Mouzaki, 2013).

Considering the terminology, the term *error* has been used in accuracy-oriented literature (Di Filippo et al., 2005; Ehri, 1991; Geva & Siegel, 2000; Gilabert, 2007; Kormos, 2000) and it implies a dichotomical distinction between correct and incorrect answer. This does not seem to cover sufficiently the model applied in the current study in the case of, e.g., self-corrections where speakers consciously and actively modify their product. The term *disfluency*, on the other hand, is used in the utterance fluency research to refer to surface properties of fluency and it is considered separate from accuracy (e.g., Bosker et al., 2013; Kahng, 2014). In the current paper, the term *inaccuracy* is therefore used, both to avoid confusion with the other theories' operationalisations and, further, to imply that the categorisation used in the current paper includes both the breakdowns relating to inefficiency and to control of attention.

Very few studies have measured inaccuracies of lexical access, and mostly only the percentage of correct answers has been of interest (e.g., Akamatsu, 2008; Gholamain & Geva, 1999; Salmi, 2008). Operationalisation or theorisation of inaccuracies in lexical access is nearly non-existent, and detailed analyses are rarely considered. One study attempting to categorise the inaccuracies in lexical access is Kaukonen and Lanu (2005), in which the authors analysed both weak and normal readers' word reading errors in L1 Finnish. They found the most frequent error types to be guessing, difficulties in pronunciation, word recognition problems, and speed–accuracy trade-off. These error types were found to differentiate the weak and normal readers significantly. Based on these results, the authors hypothesised a model for inaccuracies stemming from the undeveloped automaticity, working memory restrictions, lexical restrictions, motivation problems, and problems in maintaining alertness. Danielsson (2003) conducted a similar study in L1 Swedish, explaining the reasons for errors in word reading mainly on orthographic or linguistic principles (see also Geva & Siegel, 2000). These studies have not, however, considered the categorisations in relation to processing limitations. According to Segalowitz's cognitive fluency model (2010), the errors might be interpreted as stemming from either the inefficient language skills or from the attention-related processes. Next, I will discuss how these might be realised in lexical access.

1.2.1 Inefficient processing

First, cognitive fluency entails efficient, i.e., both fast and accurate processing. For example, to be a fluent reader, most words in the text have to be recognised as wholes (Ehri, 1991). If the language skills are limited, the speaker may have to choose between speed and accuracy, which leads to trade-off (Grabe, 2009, p. 292; Kame'enui & Simmons, 2001; Walczyk et al., 2004). Favouring speed over accuracy may lead to, for example, *skipping* items, whether by accident or for buying time (cf., Marian, Blumenfeld, Mizrahi, Kania & Cordes, 2013). When aspiring to read quickly, the readers may also revert to *guessing* words, especially when encountering low-frequency words (see Balota et al., 2006, for an overview). The low-frequency words are often read as more frequent items of similar length and overall graphemic structure, i.e., visual form (Balota, Yap & Cortese, 2006; Broadbent, 1967; Danielsson, 2003). Similarly, Ellis (2002) notes that in listening, more common words are perceived correctly more quickly, and incorrect responses usually stem from a small set of relatively common words. Cossu, Shankweiler, Liberman, and Gugliotta (1995) did not find visual form affecting misreadings of individual letters as much as phonological interference (b–d and d–t but not u–n), but the case might be different for whole words (e.g., word superiority effect, see Balota et al., 2006). Gollan and her colleagues (2011) found that word frequency was more closely related to word recognition than to naming speed (example of word retrieval), which involves only highly familiar words. Inefficiency thus manifests itself most likely in different ways in word recognition and retrieval tasks, and in the current setting, it was possible to compare if this was the case.

1.2.2 Attentional processes

Attention-control may also induce breakdowns of fluency, e.g., in the form of repairs (Skehan, 2003). *Correcting and monitoring* one's own production disrupts the flow of speech, but also requires cognitive resources. For the beginning language learners, there may not be enough resources to allocate to these higher processes (cf., Kahng, 2014; Kormos, 2000). In the previous literature, self-corrections have been hypothesised to be more prominent amongst the more proficient language learners but empirical findings are scarce (Lennon, 1990; Segalowitz, 2010; cf., Danielsson, 2003). In Kaukonen and Lanu's (2005) data, the number of self-corrections did not decrease from second to third grade for either the normal or weak readers. Freed (2000) found that students with more advanced speaking skills (stay-abroad experience) attempted to correct their expressions and to reformulate their speech to a greater extent than their less advanced peers, thus producing more false starts. Repetitions and self-corrections, further, did not differentiate between the at-home and abroad groups. For *repetitions*, Peltonen and Lintunen (2016) found that these were more connected with personal speaking styles and strategies than inadequate language skills in their L2 speech data. Repetitions showed very much within-group variation and their use did not mostly differentiate between proficiency groups (see also Bosker et al., 2013). This implies that speakers may use repetitions in a functional manner, to avoid excessively long silences, to keep the speech-turn, and to minimise disruptions in the speech (Peltonen & Lintunen, 2016). The use of repetitions, furthermore, did not differentiate the ADHD and normal control groups in a sentence production task (Engelhardt et al., 2010). In the study by Kahng (2014), the repetitions correlated only weakly with speaking scores and corrections did not correlate at all, although the L2 speakers used these repairs more than the L1 speakers did. The role of the repairs in language proficiency remains thus quite unclear. In the current study, it is hypothesised that these might be related to more advanced language skills, when there are more resources to apply them.

Attention-control entails, furthermore, *inhibition or suppression of irrelevant information*. The difficulty in inhibiting irrelevant information is very often illustrated by the Stroop task (Stroop, 1935), where participants are instructed to name the colour of the print instead of reading a colour word. Naming times are longer when the colour and word are incongruent (e.g., word "RED" written in blue ink), and this indicates that reading of the colour name has to be suppressed in order to be able to name the ink (for a review, see MacLeod, 1991). It has been shown that, with better language skills, it is in fact more difficult to inhibit irrelevant material. Automatic activation of the word meaning is stronger in L1 than in L2, for example, and suppressing this activation requires cognitive resources (Marian et al., 2013; see also Gernsbacher, 1993). Favreau and Segalowitz (1983) showed that less-skilled L2 users benefited from a priming effect (e.g., were quicker to recognise the word "apple" when preceded by "fruit" rather than by "furniture") only when there was enough time to consciously process the material (1150ms). More skilled L2 users showed a priming effect even with shorter exposure times (200ms). This illustrates how automatic processing, and the lack thereof, depends on the language proficiency, and this affects also the inaccuracies that may be produced.

The lexical access tasks used in the current study (single word recognition and word retrieval) lack the contextual clues for priming to happen, but difficulties in suppressing irrelevant material may result in inaccuracies stemming from successive words in the list format. Eye-tracking studies offer support for the view that automatization makes inhibiting the irrelevant material more difficult. Fast readers are found to pre-process words just outside of the fixation point, and this has been shown to influence processing of the fixated word (Pollatsek, Rayner & Balota, 1986; Simola, Holmqvist & Lindgren, 2009; shown also in listening tasks: Ellis, 2002). Jones, Obregón, Kelly, and Branigan (2008) found that even in a naming speed task (RAN), there was sensitivity to some visual and phonological information from material in the parafoveal preview (i.e., material not yet focused on). This is an indication of the automatic connections to lexical storage, which are activated without conscious effort, but require higher levels of language proficiency. Therefore, it is hypothesised here that the beginning language learners do not encounter inhibition difficulties. In reading a word list and naming items, some pre-processed information from the visual field may influence the currently produced items of the more advanced L2 learners; however, they might follow a u-shape thus that the most advanced students have enough resources for suppressing to happen appropriately.

Switching between different tasks or mental sets also requires attentional capacity. This has been studied, e.g., with number-letter task where participants have to decide whether, in a number-letter combination (e.g., 7G), the given number is even or odd, or the given letter is a consonant or vowel, depending on the combination's location on a screen (Miyake et al., 2000). Shifting, when including two different languages, has been proposed to be easier from L1 to L2 than vice versa (cf., Meuter & Allport, 1999). This is because L1 needs to be heavily suppressed to be able to perform in the less automatized L2, and Segalowitz and Frenkiel-Fishman (2005) report that there is more shift-cost when switching from L2 to L1 than vice versa. As was proposed with inhibition, it may be that it is more difficult to suppress a previously activated category with more automatized language skills than in a language that is still less developed and with lower activation level. The current study explored the effects of attention-control in lexical access tasks not only between different-level learners, but also in both L1 and L2 to find possible differences between more and less automatized languages.

1.3 Aims of the present study

The current study concentrated on a categorisation of inaccuracies in lexical access. The categories were based on processing limitations in language production. The sources of inaccuracies were considered to be *inefficient processing* and *control of attention*. The hypothesis was that if lexical access is not automatized, there are no resources for attention-control. Only when proficiency develops, can cognitive resources be allocated to monitoring own production. The role of proficiency was examined, first, by the types of inaccuracies between the L2 learners of three grade levels. Second, the inaccuracies between the L1 and L2 results were compared within the grades to see if the number and the type of the inaccuracies differed according to more or less proficient languages.

2 Data and method

The data were gathered in an experimental, cross-sectional setting, with participants from three different grades as proxies for L2 proficiency levels. To measure fluency of lexical access, two standardised psycholinguistic tasks were used, originally designed for L1 diagnosis. Here, they were conducted in L2 to study their interaction with L2 proficiency in an experimental manner. In contrast to the usual measurement of speed and accuracy, here a more exploratory categorisation of inaccuracies was used. Furthermore, fluency of L1 lexical access was measured as well for comparison data. The data were analysed quantitatively by counting the number and proportions of different inaccuracies and comparing them between the grades. The aim was to see if there were differences in the occurrences of the types of inaccuracies between the proficiency levels. In addition, χ^2 tests were used to examine if any of the possible differences between the proficiency groups were statistically significant. The inaccuracies were further explored comparing the performance in L2 lexical access tasks to the similar ones in L1, to see if there were differences in the less and more proficient languages within the grade levels.

2.1 Participants

The participants (N=563, 53% females) were Finnish schoolchildren who studied English as a foreign language. There were three age groups: *Grade 4* (age = 10, n = 192), *Grade 8* (age = 14, n = 186), and *Grade 11* (school year 11, age = 17, n = 185). All the participants had learned English since Grade 3, and the grade levels were used as an approximation of the L2 level (as explained further in Olkkonen et al., submitted). As Finnish orthography is very transparent, children are mostly very fluent readers of L1 by Grade 4 (see e.g., Eklund, Torppa, Aro, Leppänen & Lyytinen, 2015). The data were gathered in DIALUKI project in Finland from 2010 to 2011. The students came from 37 different schools around Finland, in both rural and urban areas. The testing was conducted during the school hours and consents to participate in the study were obtained from the counties, the schools, the parents, and the children themselves.

2.2 Procedure and materials

The lexical access tasks were a part of a cognitive test battery (appr. 45 minutes), that was conducted individually in a quiet room by trained assistants with a *Cognitive Workshop* software (the Finnish version developed by the University of Dundee and the Jyväskylä Longitudinal Study of Dyslexia, see Lyytinen et al., 2004; the English versions by DIALUKI project). Two lexical access tasks were used: one word retrieval (RAS) and one word recognition (Word List), with similar tasks in both L1 and L2. The order of administering the L1 and L2 versions was counterbalanced to minimise the language order effect.

Rapid Alternating Stimulus (RAS) task was used for measuring naming speed and retrieval of familiar vocabulary (Wolf, 1986). In L1 for all groups and in L2 for Grade 8 and 11, the matrix consisted of 50 units, including numbers (2, 6, 7, and 9), colours (black, blue, green, red, and yellow), and letters (A, E, S, and T), arranged in five rows in a semi-random order (from Ahonen, Tuovinen & Leppäsaari, 2003). For Grade 4, the task was slightly easier and shorter in L2 and it consisted of 30 units, with numbers (2, 6, 7, and 9), colours (black, blue, green, red, and yellow), and familiar pictures (pencil, car, fish, and house). Picture naming was considered to be more automatized than foreign language letter naming at this level (see Denckla & Cutting, 1999). In comparison to simple RAN, which uses items from only one semantic category, RAS requires shifting between different semantic categories that makes it a more challenging task (see Wolf, 1986). Before the task, the assistant checked that the child knew the names of all the used items. Participants were asked to name the items aloud as fast and as accurately as they could, and the time to conduct the task was measured by a stopwatch. The task was conducted individually for each student and recorded to the computer with a microphone-headset for later analyses.

Word List reading. The children were given a printed word list of 105 words, with the words arranged in three columns, to be read from top to bottom. In the beginning of the list, the words were short and familiar, gradually becoming longer and more complex. Children were instructed to read aloud as many words as they could in 60 seconds' time, as fast and as accurately as possible. The task was conducted individually, in both L1 and L2. The task originated from the standardised Lukilasse reading test battery for L1 Finnish in primary school (Häyrinen, Serenius-Sirve & Korkman, 1999). The English word list was designed similarly in DIALUKI project by sampling words from a frequency list[1]. Even though the number of words was equal, the lists differed in the number of syllables (Finnish 379 syllables; English 183 syllables). This was due to affixation and complex inflectional morphology of the Finnish language: the longest words to be read in L1 were 22 letters long (e.g., *prosessikirjoittaminen* 'process writing' versus L2 English ten letters in *experience*). This made the Finnish word list quite challenging; however, as the phoneme-grapheme-connections are very reliable and the orthography is transparent, the decoding of the words is quite easy even for young readers (Aro & Wimmer, 2003).

2.3 Classification of inaccuracies in lexical access

To assess the question of the types of inaccuracies in relation to cognitive fluency, a classification of the inaccuracies was devised, based on the Segalowitz's (2000, 2010) fluency construct. The inaccuracies were proposed to stem from two sources, inefficiency and attention-control. First, the inaccuracies that relate to inefficiency were hypothesised to be visible with lower-proficiency L2 levels, and to decrease when proficiency increases. Second, the control of attention was proposed to show in more proficient language-learners who have more resources for the controlled processes, but may also encounter more involuntary activation of irrelevant material. All the responses in the RAS and Word List tasks were analysed according to the classifications described below (summarised in Table 1.). The inaccuracies were coded from the sound files that were recorded at the testing situation. The number of participants included in each analysis differed slightly between the tasks due to technical problems in the recordings. One item could include several inaccuracies (e.g., correcting a skipped item). As the classifications were exploratory and without clear correct / incorrect dichotomy, 10% of the data from each grade were double-checked from the sound files by a second rater, informed on the original assessment criteria. The overall agreement rate on accuracy, and place and type of inaccuracies was 95.6%.

a) Inefficient processing
   The inefficiency inaccuracies were proposed to stem from non-automatic lexical access, where processing capacity was required for recognising and retrieving words. This was assumed to relate to lower proficiency of L2 skills: i.e., to be found more for the beginning language learners and to decrease with more proficiency. In addition to the possible categories discussed in Chapter 1.2, errors within the same category and pronunciation difficulties were included.

1. category *Guessing (GUE)*. Fast performance requires sight-word reading (Ehri, 1991), and when word recognition is not automatized this can lead to guessing words. In the RAS task, this category included visually similar items, for example the letter '*A*' mistakenly named as the number '*4*'. In word recognition, guessing often follows a frequency-based bias, which results in inaccuracies that are graphemically similar, but more frequent ones than the target words (Broadbent, 1967; Balota et al., 2006). This category of inaccuracies was hypothesised to decrease as proficiency increased. The category included only real words (Example 1) and possible, readily understandable neologisms (Example 2); nonwords were not included (similarly as Danielsson, 2003).

Example 1

Finnish	heilahdella 'swinging' > heilahdus 'a swing'
	turkikset 'furs' > turkkilaiset 'Turkish'
	Ilmarinen (character in Finnish mythology) > ilmainen 'free', imarrella 'to flatter'
	haluttaisiin 'would be wanted' > hautajaisiin 'to the funeral'
English	probable > probably, course > of course
	place > please, move > movie, find > Finland

Example 2

Finnish	professori 'a professor' > professoida 'to (act like a) professor'
	kierrätyskeskus 'recycling centre' > kierrähdyskeskus 'revolving centre'
English	particular > practicular

2. category *Skipping (SKIP)*. Inefficiency of lexical access was further proposed to be visible in skipping individual words in Word List or items in RAS, whether purposefully or accidentally, when aspiring for speed (cf., Marian et al., 2013).
3. category *Same-category (SCAT)*. Inefficient lexical access in RAS task was interpreted to show in inaccuracies within the same category (e.g., blue > green). The requirement for speed was proposed to cause inaccuracies even with very automatized material. The results were controlled for the possible colour-blindness.
4. category *Pronunciation problems (PRON)*. This category included miscellaneous problems in pronouncing items, e.g., reading English words according to Finnish grapheme-to-phoneme rules, and otherwise unclassifiable instances of pronunciation problems.

b) Attentional processes
Inaccuracies relating to control of attention were hypothesised to increase with proficiency, relating to two trends. First, with higher proficiency, there are more resources available for monitoring own speech as the lower-level language processes are becoming automatized. On the other hand, with higher proficiency, there is also a possibility of automatic activation of lexical material that is involuntary and can disturb the production.

5. category *Self-correction (SCOR)*. Correcting one's own production was considered to require processing capacity and to increase with increasing proficiency of language skills (Walczyk et al., 2004; Segalowitz, 2010; also Schmidt, 1992). No further distinctions were drawn on these, and whether the correction was successful or not was not included in the analyses (cf., Kormos, 2000).
6. category *Repetition (REP)*. Repeating items is included in the repair phenomena of speech (Skehan, 2003), and therefore, it was considered here to be a part of conscious monitoring of speech. Repetitions are proposed, furthermore, to be possible strategic devices and an intentional way of buying time during speech (Peltonen & Lintunen, 2016). In the current study, this category only included repeating whole items, whereas partial repetitions were included in pronunciation problems (cf., De Jong et al., 2015).
7. *Inhibition difficulties (INH)*. As the processing of words begins slightly before the gaze is fixated on them (Jones et al., 2008; Simola et al., 2009), succeeding material can affect performance especially in time-pressure situations. This is shown in the Example 3 where a plural ending *–eet* from word number 43 is also copied to the word number 42. This was hypothesised to be an indication of more fluent reading, and involuntary parafoveal word processing.

Example 3

Finnish 42. kyynel > kyynel/eet 'tear > tear/s'
43. pyyhk/eet 'towels'

8. *Shifting difficulties (SHIF)*. In the RAS task, the difficulty of suppressing an activated category was proposed to cause inaccuracies in category shifting. An example of category shifting difficulty is when the number *six* is followed by the letter '*S*', and is, therefore, named to be the number *seven*. The strength of activation is proposed to depend on the proficiency level, so that

with more automatic language skills the activation is also stronger (Segalowitz & Frenkiel-Fishman, 2005).

Table 1. Summary of the sub-categories of inaccuracies in lexical access and their descriptions.

Category	Description
Inefficiency	
1. Guessing (GUES)	producing wrong responses that are graphemically similar to the target words, frequency-based (Broadbent, 1967; Balota et al., 2006)
2. Skipping (SKIP)	skipping items either accidentally or strategically (Marian et al., 2013)
3. Same-category (SCAT) * only in RAS task	producing wrong responses that are within the same category, such as naming 6 as 7, A as E, or green as blue
4. Pronunciation problems (PRON)	stammering when encountering difficult phonemes or combinations
Attention	
5. Self-correction (SCOR)	correcting one's own speech requires attention and cognitive resources (Segalowitz, 2010)
6. Repetition (REP)	repeating items, subject to individual differences and strategic use (e.g., Kahng, 2014; Peltonen & Lintunen, 2016)
7. Inhibition difficulties (INH)	difficulty with suppressing irrelevant information are proposed to be connected with more advanced skills as the automatic activation is stronger (Marian et al., 2013; Jones et al., 2008)
8. Set-shifting (SHIF) * only in RAS task	set-shifting is more difficult with stronger activation and may be easier with less proficiency; e.g., switching from activated letters to numbers (Segalowitz & Frenkiel-Fishman, 2005)

3 Results

Overall, the accuracy in both of the lexical access tasks was very high. Accuracy percentages for English L2 RAS were 96% (Grade 4), 93% (Grade 8), and 97% (Grade 11). For Finnish L1 RAS, the accuracy percentages were 96% (Grade 4), 98% (Grade 8), and 99% (Grade 11). The accuracy percentages for English L2 Word List were 88% (Grade 4), 95% (Grade 8), and 97% (Grade 11). To compare, the Finnish L1 the accuracy percentages for Word List were 89% (Grade 4), 94% (Grade 8), and 96% (Grade 11; for further discussion on the overall accuracy in the tasks, see Olkkonen et al., submitted). The number of inaccuracies was, therefore, small, as most of the items named or read were correct and fluent. Nonetheless, the distributions of inaccuracies can still offer some insights into differences between the proficiency levels, even though the distributions were mostly too skewed for rigorous statistical analyses.

Thus, first relative proportions of inaccuracies were counted to compare the performance *between* the grades, in both languages. The number of items produced was different in both L2 RAS, where the number of items differed between Grade 4 (30 items) and the higher grades (50 items), as well as in L1 and L2 Word Lists, where the number of words read varied for each individual because of the time limit. Therefore, the proficiency levels could not be directly compared. The proportions were counted by dividing the number of each inaccuracy by the total number of words read / named (possibilities for inaccuracies) for each subgroup separately. Thus, a percentage of potential occurrences was achieved (see also Protopapas et al., 2013). Second, χ^2 tests were performed to see if the proficiency levels yielded any significant differences in the use of different inaccuracies, comparing the number of children who made no errors vs. children with at least one error of a certain type within the proficiency level. Third, the frequencies and distributions of the different inaccuracy types were counted for each grade, especially to compare the performance in the L2 tasks to the L1 versions *within* the grades. This was considered to reveal possible tendencies between the more and less proficient languages. The control measures, i.e., L1 fluency, to which the L2 results were compared to, are presented in the tables first, followed by the L2 measures.

3.1 Differences in the inaccuracies between the grades

To examine the differences between proficiency levels as regarding the grades, the proportions of inaccuracies in each proficiency level relative to the total number of occasions (words or items) were counted. The largest category of inaccuracies for all groups was the pronunciation problems, and it is presented in a separate figure (Figure 1.) to show the other categories more clearly (in Figures 2.–5.). In RAS L1, there were no differences between the groups in pronunciation difficulties. In the RAS L2, however, Grade 8 results were highly affected by the difficulty in pronouncing the letters A [eɪ] and E [i:] (49% of all their pronunciation problems). The pronunciations for the Finnish vowels are A [ɑ], E [e], and I [i], which causes them to be easily confused in English. The matrix in Grade 4 did not include letters, which is why the results of Grade 4 RAS L2 are not directly comparable to the other grades' results, but the amount of pronunciation problems decreased from Grade 8 to 11. In the Word List, the number of pronunciation problems decreased steadily by proficiency in both language, which indicates that the inefficiency of lexical access decreased with developing skills.

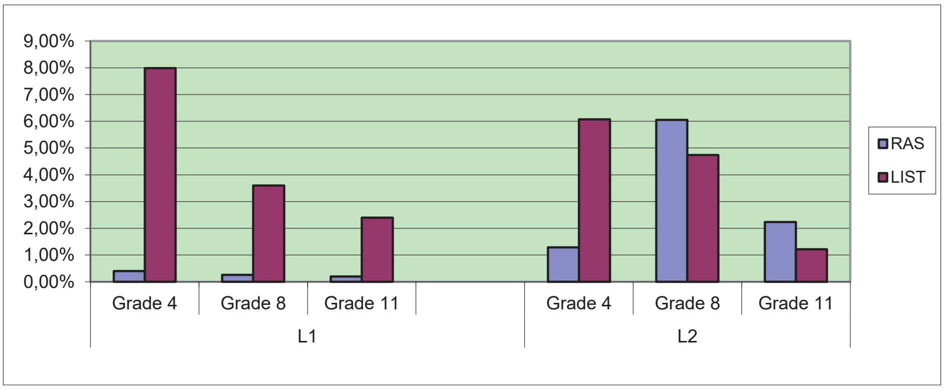

Figure 1. Distributions of relative proportion of pronunciation problems (total number of inaccuracies by category divided by total number of words) by Grades in RAS and Word List tasks.

For the other inaccuracies, in the L1 RAS task, the numbers in both inefficiency and attention-control categories decreased with proficiency (Figure 2.); i.e., most types of inaccuracies decreased with automatization of word retrieval. More skipped and repeated items were found for Grade 4 than for the other grades. In the same-category inaccuracies (SCAT) the groups were indistinguishable. Overall, in RAS L1, even for Grade 4, the accuracy was very high and differences between proficiency levels were quite minimal.

In L2 RAS, the trends were not as clear, and the results of Grade 4 are not directly comparable to the other results. However, the inaccuracies in the inefficiency categories, skipping (SKIP) and same-category errors (SCAT), seemed to decrease with proficiency (Figure 3.). On the other hand, for the attention-control categories, the self-corrections followed an inverted U-shape, as in Grade 8 these were found more often than in the other groups. In repetitions, there were almost no differences between the groups. Guessing, inhibition difficulties, and shifting were almost non-existent in L2.

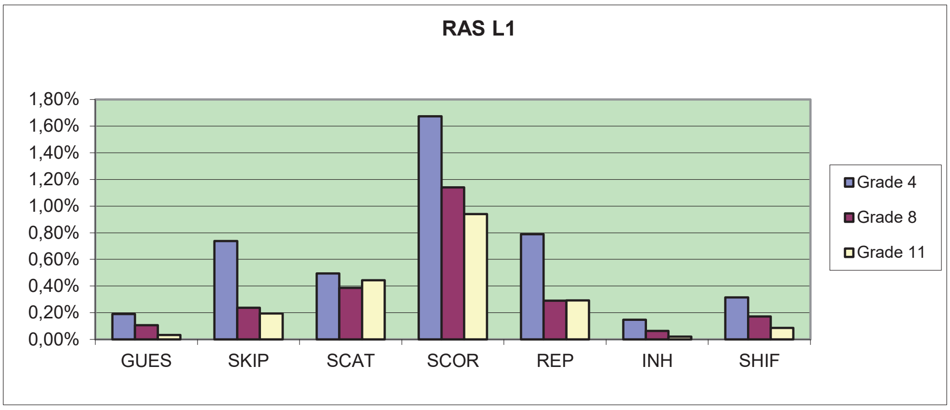

Figure 2. Distributions of relative proportions (the total number of inaccuracies by category divided by total number of words) by Grade in L1 RAS task.

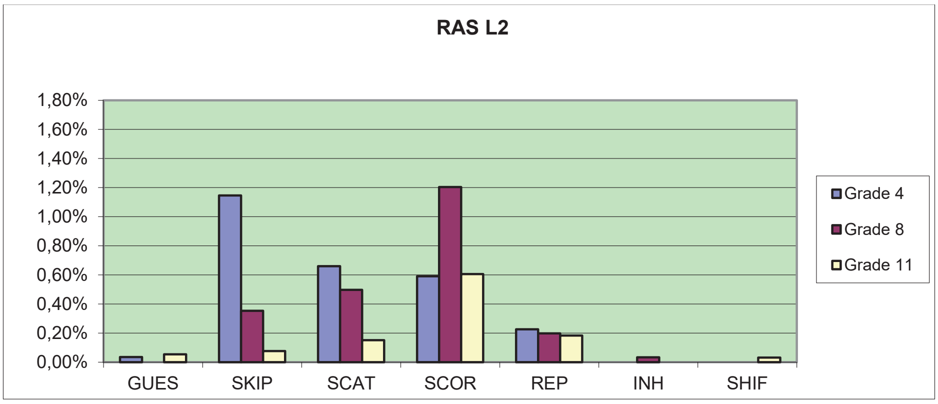

Figure 3. Distributions of relative proportions (the total number of inaccuracies by category divided by total number of words) by Grade in L2 RAS.
Note. GUES = Guessing. SKIP = Skipping. SCAT = Same-category. SCOR = Self-corrections. REP = Repetition. INH = Inhibition. SHIF = Category shifting.

In the Word List task, the inefficient processing in the form of guessing (GUES) decreased by proficiency in both L1 and L2 (Figures 4. and 5.). Skipping was found almost exclusively in L2 and its use decreased by proficiency, but it was used very rarely. For the attentional processes, the number of self-corrections followed slightly an inverted U-shape as in the RAS task: increasing from Grade 4 to Grade 8, and then decreasing again. For repetitions, there were no differences between the grades in L1, and in L2 their number decreased. The number of inhibition difficulties decreased by proficiency and were found almost exclusively in the L1 task.

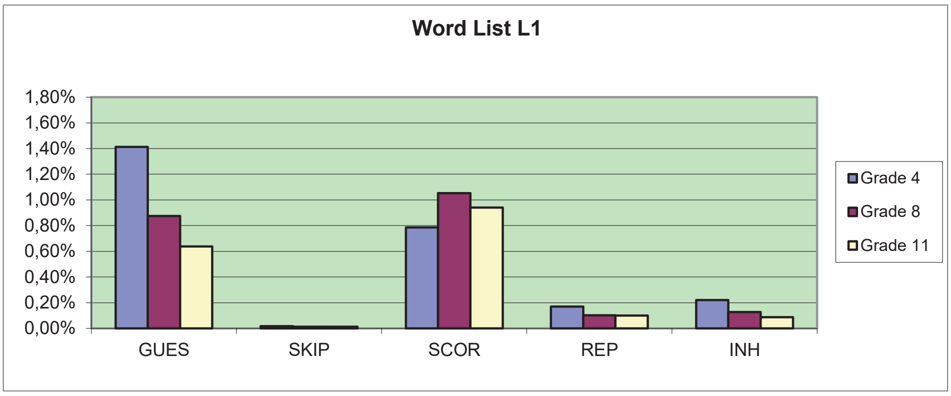

Figure 4. Distributions of relative proportions (the total number of inaccuracies by category divided by the total number of words) by Grade in L1 Word List.

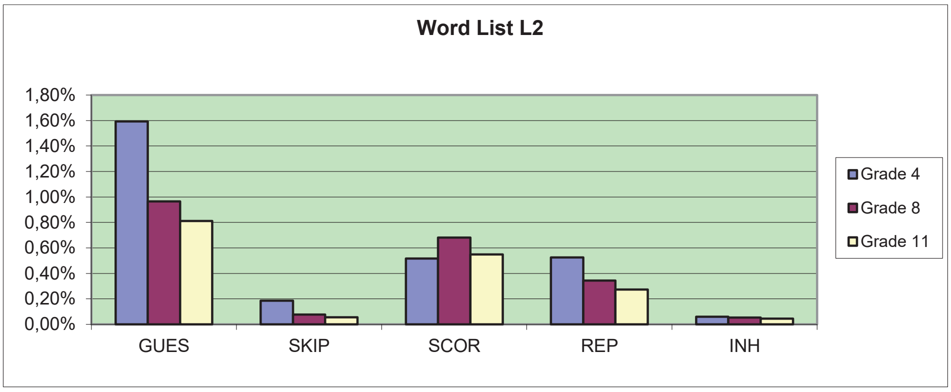

Figure 5. Distributions of relative proportions (the total number of inaccuracies by category divided by the total number of words) by Grade in L2 Word List.
Note. GUES = Guessing. SKIP = Skipping. SCOR = Self-corrections. REP = Repetition. INH = Inhibition.

Second, to test if any of the apparent differences between the grades were statistically significant, χ^2 scores were counted for each inaccuracy category. These compared the proficiency levels in relation to the proportions of children in each grade who used a particular type of inaccuracy. In RAS L2, Grade 4 was left out of the analysis due to the task differences. Grade 8 had more inaccuracies than expected, compared to Grade 11 performance, in self-corrections (χ^2 (1, 427) = 8.44, $p < .01$, adjusted standardized residual = 2.9), in same-category errors (χ^2 (1, 427) = 10.76, $p < .001$, adjusted standardized residual = 3.3), and in pronunciation problems (χ^2 (1, 427) = 8.45, $p < .01$, adjusted standardized residual = 2.9). Other inaccuracy types did not yield significant effects. To compare to RAS L1 results, χ^2 tests showed modestly significant differences between the proficiency levels for self-corrections and repetitions. Grade 4 children differed from the other two grades, as they corrected themselves more than expected (χ^2 (2, 638) = 8.03, $p < .05$, adjusted standardized residual = 2.8) and repeated more items (χ^2 (2, 638) = 6.93, $p < .05$, adjusted standardized residual = 2.5).

In the Word List, the χ^2 test showed significant effects between the groups for self-corrections. There were fewer errors than expected for Grade 4 in both L1 (χ^2 (2, 638) = 22.09, $p < .001$, adjusted standardized residual = -4.6), and L2 (χ^2 (2, 638) = 17.03, p < .001, adjusted standardized residual = -4.1). Furthermore, a significant effect was found for Grade 4 in L2 pronunciation problems, for more than one inaccuracy (χ^2 (4, 638) = 14.47, $p < .01$, adjusted standardized residual = 2.4). To compare to L1 results, more pronunciation problems were found for Grade 4 (χ^2 (2, 638) = 25.22, $p < .001$, adjusted standardized residual = 4.5), and Grade 11 skipped significantly more in L1 than expected (χ^2 (2, 638) = 7.28, $p < .05$, adjusted standardized residual = 2.7). Other inaccuracy types did not yield significant effects.

One additional note is that the patterns from χ^2 tests showed some support for the proportional analyses (see Chapter 3.1) in that all the groups seemed to self-correct less in L2 than in L1. In Word List, self-correcting in L2 was used in Grade 4 by 31% of the children, in Grade 8 by 49%, and in Grade 11 by 49%, as compared to L1 (4: 39%, 8: 57%, and 11: 61%). Similar pattern was found for inhibition difficulties (4: L2 12% vs. L1 19%, 8: L2 15% vs. L1 19%, and 11: L2 21% vs. L1 23%). In RAS, Grades 8 and 11 had fewer difficulties with set-shifting in L2 than in L1 (13% vs. 17% and 16% vs. 19%). In Grade 11, a smaller proportion of students corrected themselves in L2 than in L1 (36% vs. 45%), as well as skipped items (18% vs. 23%).

3.2 Differences in the inaccuracies between the L1 and L2

The frequencies and percentages of different types of inaccuracies across the grades are presented in Tables 2 and 3. The comparisons offer information especially on the group profiles in their performance in L2 as compared to L1, regarding the proportions of the two main categories, i.e., inefficiency-related inaccuracies and the attention-control inaccuracies. Overall, the results were quite similar between the groups regarding the most common types of inaccuracies, especially for the Grades 8 and 11. In RAS L1, the self-corrections (SCOR) constituted the largest proportion of inaccuracies in all grades (4: 35%, 8: 43%, and 11: 43%; see Table 2.). However, Grade 4 had the repetitions (REP, 17%) as the second most common type of inaccuracy, whereas the Grades 8 and 11 had same-category inaccuracies (SCAT, 15%). In RAS L2, the largest group of inaccuracies was the miscellaneous pronunciation problems (PRON) for all groups (4: 33%, 8:

73%, and 11: 67%). The second most common type of inaccuracy in RAS L2 was skipping (SKIP) for Grade 4 and self-corrections (SCOR) for Grades 8 and 11. To summarise, for all the groups, the inaccuracies in L1 belonged mostly to the attention-related categories (SCOR, REP), whereas in L2, the inaccuracies related mostly to inefficiency (PRON, SKIP). However, for the more proficient students, attention-related inaccuracies were the second common type also in L2 (SCOR). A further point to notice is that the last two types of attention-related inaccuracies (inhibition difficulties and shifting), that were rare even in L1, were practically non-existent in L2.

Table 2. Frequencies and percentages of different inaccuracies by language and grade in RAS.

	L1						L2					
Grade	**4**		**8**		**11**		**4**		**8**		**11**	
Inefficiency	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
GUES	18	4.0	10	4.0	3	1.5	2	0.9	0	0	5	1.6
SKIP	70	15.5	22	8.9	18	8.8	66	**28.9**	32	4.2	7	2.3
SCAT	47	10.4	36	**14.6**	41	**20.1**	38	16.7	45	6.0	14	4.5
PRON	38	8.4	24	9.7	18	8.8	74	**32.5**	548	**72.6**	207	**66.8**
Total %		38.3		37.2		39.2		79.0		82.8		75.2
Attention	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
SCOR	159	**35.3**	106	**42.9**	87	**42.6**	34	14.9	109	**14.4**	56	**18.1**
REP	75	**16.6**	27	10.9	27	13.2	13	5.7	18	2.4	17	5.5
INH	14	3.1	6	2.4	2	1.0	0	0	3	0.4	1	0.3
SHIF	30	6.7	16	6.5	8	3.9	1	0.4	0	0	3	1.0
Total %		61.7		62.7		60.7		21.0		17.2		24.9

Note. GUES = Guessing. SKIP = Skipping. SCAT = Same-category. PRON = Pronunciation problems. SCOR = Self-corrections. REP = Repetition. INH = Inhibition. SHIF = Category shifting. Total number of items in the task: L1 Grade 4: 9,500, 8: 9,250, 11: 5,760; L2 Grade 4: 9,300, 8: 9,050, 11: 9,250.

For the Word List, the pronunciation problems constituted the largest percentage of inaccuracies in both L1 (4: 75%, 8: 62%, and 11: 57%) and L2 (68%, 69%, and 41%, respectively; see Table 3.). The most difficult words in L1 Finnish were either the very long ones (*nimikkoluokka* 'dedicated class'), or the ones containing combinations such as diphthongs *kiulu* 'pail' > *kuilu* 'gorge' or liquid consonants *broileri* 'broiler' > *bloireri*. In L2 English, the problematic words often included phonemes that are not part of Finnish phonotactics, such as [ʧ] or [ð]: much > [mɑks], church > [kurs]. The second most common type of inaccuracy in the Word List for Grade 4 was guessing (GUES) in both L1 (13%) and L2 (18%), as well as in

L2 in Grades 8 (14%) and 11 (28%). In the L1 Word List, the Grades 8 and 11 used self-corrections considerably (18% and 24%, respectively). To summarise, the largest proportion of inaccuracies for all proficiency levels and in both languages resulted from inefficiency. It is notable, however, that the proportion of inaccuracies relating to attentional processes increased by grade in both L1 and L2.

Table 3. Frequencies and percentages of different inaccuracies by Grade and language in Word List.

	L1						L2					
Variable	4		8		11		4		8		11	
Inefficiency	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
GUES	167	**13.3**	137	15.2	103	15.3	188	**17.8**	163	**14.1**	145	**27.5**
SKIP	2	0.2	2	0.2	2	0.3	22	2.1	13	1.1	10	1.9
PRON	944	**75.4**	564	**62.4**	387	**57.4**	717	**67.8**	801	**69.1**	217	**41.2**
Total %		88.9		77.8		73.0		87.7		84.3		70.6
Attention	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
SCOR	93	7.4	165	**18.3**	152	**22.6**	61	5.8	115	9.9	98	18.6
REP	20	1.6	16	1.8	16	2.4	62	5.9	58	5.0	49	9.3
INH	26	2.1	20	2.2	14	2.1	7	0.7	9	0.8	8	1.5
Total %		11.1		22.3		27.1		12.4		15.7		29.4

Note. GUES = Guessing. SKIP = Skipping. PRON = Pronunciation problems. SCOR = Self-corrections. REP = Repetition. INH = Inhibition. Total number of items in the task: Grade 4 L1: 11,828, 8 L1: 15,675, 11 L1: 16,148; Grade 4 L2: 11,809, 8 L2: 16,902, 11 L2: 17,865.

4 Discussion

The present study aimed to see whether the differences in the types of inaccuracies in lexical access could be attributed to the L2 proficiency levels. The inaccuracies in lexical access were proposed to stem from two sources: inefficient processing (cf., Walczyk et al., 2004; Grabe, 2009, p. 292) and attention-control (Kormos, 1999, 2000; Gilabert, 2007; Miyake et al., 2000; Segalowitz, 2000, 2010; Jones et al., 2008). It was proposed that these two sources of inaccuracies would behave differently in relation to proficiency: if lexical access is not automatized, there are no resources for attention-control. Therefore, the first category was proposed to decrease with proficiency, whereas the second was hypothesised to be more stable or to increase. For measuring lexical access, one word retrieval (RAS) and one word recognition (Word List reading) task were used. Finnish children learning English from Grade 4, Grade 8, and Grade 11 participated in the study. These grade levels were used as estimates of L2 proficiency. Overall, the findings suggest that the types of inaccuracies were connected to proficiency to some extent: in the L1 tasks and for the older students, the inaccuracies relating to attention-control accounted for the largest proportion of the inaccuracies (self-corrections). However, in L2 and for the Grade 4, the difficulties seemed to stem more from the lack of efficiency (guessing, skipping, and pronunciation problems). Statistically significant differences were found especially for self-corrections (attention-control) and pronunciation problems (inefficiency). These inaccuracies, therefore, were shown to be the most robust indicators of their respective categories. The other categories mostly offered interesting possibilities for future studies, as the number of occurrences was quite limited and the tasks were not designed to measure these categories of inaccuracies.

4.1 Inefficient processing

The inaccuracies connected with the inefficient processing constituted the largest proportion of inaccuracies in the L2 tasks, especially for the youngest group (Grade 4). The number of inefficiency inaccuracies, as well their proportion of total of inaccuracies, was found to decrease with proficiency. These results are in line with the original hypothesis that these types of inaccuracies are connected with lower language proficiency. A few points on the sub-categories are worth noting. First, the miscellaneous pronunciation problems were the largest category of inaccuracies in the L2 tasks as well as in L1 Word List, and proficiency level significantly affected their amount in both languages. Second, guessing in Word List mostly resulted in real words: either conjugated or in some other way modified forms, as in the study by Danielsson (2003). Completely different words were mostly more familiar or frequent items (cf., Ellis, 2002); even the neologisms graphemically resembled the target words. In the RAS, guessing based on form happened very rarely, which is in accord with, for instance, Gollan et al. (2011). Naming the numbers and colours was so automatic that guessing was rarely needed. Comparing the relative proportions, guessing in RAS was more frequent in L1 than L2, and this could indicate too fast a performance, with no time to notice the errors or to spare for correction.

Regarding skipping items, when looking at the relative proportions, in the RAS task children in Grade 4 and 11 seem to have skipped more items in L1 than in L2. In the other instances, in Grade 8 and the Word List task, skipping was found more in L2. These seemingly contradictory results may be due to different reasons for skipping. In the easier tasks or for the more proficient students, skipping may have resulted from the (too) fast performance, but in the more difficult task and for the beginning language learners, there may have been more strategic reasons for skipping difficult-looking items. More support for this interpretation comes from the χ^2 results, which showed that Grade 11 skipped significantly more than expected in the L1 Word List. In the future studies, this interpretation could be examined, for example, with a stimulated recall setting (Kahng, 2014), or eye-tracking paradigm. However, the findings of Marian et al. (2013) of strategic skipping in a Stroop task, offer this hypothesis some confirmation. Furthermore, skipping, when it happened in L1, was often noticed and corrected; in L2, this was not the case. This is in line with the overall processing limitations hypothesis with more resources available in L1 to monitor errors and correct one's own speech (e.g., Kormos, 2000).

4.2 Attentional processes

The inaccuracies connected with the control of attention were found to constitute the largest part of the inaccuracies in the RAS L1 task, and their proportion of the total number of inaccuracies increased with proficiency. Furthermore, they did not follow a decreasing trend, which is in accord with the initial hypothesis that attention-control requires more resources (cf., Lennon, 1990; Gilabert, 2007; Segalowitz, 2010). In the Word List task, Grade 4 self-corrected significantly less than expected in both L1 and L2, as compared to the behaviour of the more proficient students, which further offers support the monitoring resources hypothesis. In L2, the relative proportions of self-corrections increased from Grade 4 to Grade 8 in both tasks. This could indicate that in the beginning of language learning, there is limited capacity and skills to correct one's speech, and these develop only gradually. On the other hand, it has to be taken into account that the beginning language learners may also lack the L2 skills to correct themselves and also the whole monitoring system is still developing (see e.g., Diamond, 2013). However, in Grade 11 self-correcting decreased again, which may relate to one's increasing proficiency, when the need for self-correction is smaller.

For the repetitions, there were no clear differences between the groups or languages, which is in accord with the previous findings that they seem to be more related to personal style and strategic use (Engelhardt et al., 2010; Bosker et al., 2013; Peltonen & Lintunen, 2016). The only significant difference was found for RAS L1, where the children in Grade 4 repeated more than the other groups. This could be explained by the strategic use for stalling (cf., Engelhardt et al., 2010). As the Grade 4 children found it often difficult to concentrate on the tasks, it would be very interesting to address if this had any effect on the performance by comparing the distribution of inaccuracies for example along the task. This was, however, outside the current paper's scope.

The tasks used in the present study did elicit only modest amounts of inaccuracies relating to difficulties in inhibiting and set-shifting, and the hypothesised u-shape was not found. There were some tendencies, nevertheless. The number of both difficulties in inhibiting and set-shifting decreased with proficiency, and they were found almost solely in L1, in accordance with previous studies (cf., Favreau & Segalowitz, 1983; Kaukonen & Lanu, 2005; Jones et al., 2008; Marian et al., 2013). These results can be interpreted in line with the hypothesis of easier suppression of irrelevant material and lesser activation of different RAS categories in L2. This would, however, need further confirmation from tasks better designed to measure these processes (see e.g., Diamond, 2013). Nevertheless, the findings concerning attentional processes suggest that these categories do not seem very reliable indicators of proficiency (in line with, e.g., Freed, 2000; Bosker et al., 2013). Therefore, the results of the study call for caution in some of the aspects of the oral fluency assessment; e.g., self-corrections might be better interpreted as active control of the output instead of breakdowns of speech.

4.3 Theoretical implications and limitations of the study

To sum up, the current study assigned the sources of inaccuracies to processing limitations in language use (also Lennon, 2000). The limited language skills can manifest in inefficient lexical access, and when they gradually automatize, resources are freed up for the attention-requiring processes. The current finding that the more proficient student groups used self-corrections more than the less-advanced group is an example of this (also e.g., Lennon, 1990), and in some instances may relate to the strategies of the language users (Kormos, 1999). Thus, the overall proposition is that not only the frequency of the inaccuracies, but also their origin is important when studying the relationship of fluency and proficiency. The inaccuracies originating from the attentional processes seemed more related to higher proficiency, and the automatization of language processes may also lead to involuntary activation of irrelevant material (similarly in the Stroop task: Marian et al., 2013).

The results presented here, although exploratory in nature, nevertheless offer insight into the construct of fluency as a multifaceted phenomenon that should not be treated in a straightforward manner (see also, Kormos, 1999, 2000; Bosker et al., 2013; Peltonen & Lintunen, 2016). There are different trends influencing fluency in speech, especially in L2, and this study highlighted the cognitive processing aspects. One further notable point in the current data is that although L1 data was gathered more as baseline data, lexical access even in L1 was found to be far from fluent. This was especially visible in the Word List task, which also illustrated the different challenges the readers encounter in Finnish and in English. Although Finnish orthography is shallow, the long, inflected words proved to be quite difficult especially under the time-pressure of the current tasks. Therefore, the overall accuracy rates in English were in fact higher in Grade 8 and Grade 11, although a more lenient attitude of the raters in assessing L2 speech has to be considered as well. Overall, these results highlight the difficulty of using strict fluency criteria in assessment. Perceived or global fluency may be a good predictor of L2 proficiency, but when deconstructed, the connections of different phenomena are very hard to pin down (cf., Koponen & Riggenbach, 2000; Freed, 2000; Segalowitz, 2010; Bosker et al., 2013). The framework presented in the current paper is of course only one possibility and other highly interesting aspects for further study include, for example, the emotional issues such as anxiety and its influence on word retrieval and types of inaccuracies (see e.g., Housen & Kuiken, 2009).

For the limitations of the current study, first, it has to be kept in mind that these results apply only to lexical access measurement and not to longer monologue or dialogue settings. On the other hand, following Segalowitz (2010, p. 75–76; also Gholamain & Geva, 1999), lexical access was considered to offer a more pure insight into cognitive fluency, as more general planning and conversational problems were kept to a minimum. Second limitation relates to the tasks used here, as they were not designed to reveal all the types of inaccuracies studied here, and have previously been used for clinical diagnosis of dyslexia. Thus, their purpose originally has not been to distinguish normally developing children in detail, but to diagnose problems. Therefore, the tasks were often not challenging enough for the higher grades to produce enough inaccuracies for the statistical analyses, as the material was much skewed. Nonetheless, these results offer interesting tendencies and questions. Answers to them should be further explored with the help of more fine-tuned and better-suited tasks. More online-methods of measuring would also offer more detailed information on the types of inaccuracies that were only hypothesised indirectly. For instance, skipping might be studied with eye-tracking paradigm to reveal possible differences in the voluntary vs. accidental skipping, and the inhibition difficulties and involuntary activation in lexical access could be tapped with, e.g., bilingual Stroop task.

Endnote

[1] Leech, G., Rayson, R., & Wilson, A. (n.d.) Word Frequencies in Written and Spoken English: based on the British National Corpus. Retrieved October 15, 2010, from http://ucrel.lancs.ac.uk/bncfreq/lists/1_2_all_freq.txt

Acknowledgements

I wish to thank Jaana Haapasalo for her help in the additional data coding, and Lea Nieminen, Kenneth Eklund, Rob Schoonen, and Paula Salmi for their comments on the previous versions of the manuscript. Furthermore, I am grateful to the two anonymous reviewers for their input and insight.

References

Ahonen. T., Tuovinen, S., & Leppäsaari, T. (2003). Nopean sarjallisen nimeämisen testi [The test of rapid serial naming. Research reports.]. Jyväskylä, Finland: Haukkaranta school & Niilo Mäki Institute.

Akamatsu, N. (2008). The effects of training on automatization of word recognition in English as a foreign language. *Applied Psycholinguistics*, *29*(2), 175–193.

Aro, M., & Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Applied Psycholinguistics*, *24*(4), 621–635.

Balota, D. A., Yap, M. J., & Cortese, M. J. (2006). Visual word recognition: The journey from features to meaning (A travel update). In M. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed.) (pp. 285–375). NY: Elsevier.

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

Broadbent, D. E. (1967). Word-frequency effect and response bias. *Psychological Review*, *74*(1), 1–15.

Cossu, G., Shankweiler, D., Liberman, I. Y., & Gugliotta, M. (1995). Visual and phonological determinants of misreadings in a transparent orthography. *Reading and Writing: An Interdisciplinary Journal*, *7*(3), 237–256.

Council of Europe (2007). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Danielsson, K. (2003). *Beginners read aloud. High versus low linguistic levels in Swedish beginners' oral reading*. Stockholm, Sweden: Almqvist & Wiksell International.

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

Denckla, M. B., & Cutting, L. E. (1999). History and significance of Rapid Automatized Naming. *Annals of Dyslexia*, *49*, 29-42.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135–168.

Di Filippo, G., Brizzolara, D., Chilosi, A., De Luca, M., Judica, A., Pecini, C., Spinelli, D., & Zoccolotti, P. (2005). Rapid naming, not cancellation speed or articulation rate, predicts reading in an orthographically regular language (Italian). *Child Neuropsychology*, *11*(4), 349–361.

Ehri, L. (1991). Development of the ability to read words. In R. Barr, M. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (pp. 383–417). NY: Longman.

Eklund, K., Torppa, M., Aro, M., Leppänen, P. H. T., & Lyytinen, H. (2015). Literacy skill development of children with familial risk for dyslexia through grades 2, 3, and 8. *Journal of Educational Psychology*, *107*(1), 126–140.

Ellis, N. C. (2002). Frequency effects in language acquisition: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, *24*(2), 143–188.

Engelhardt P. E., Corley, M., Nigg, J. T., & Ferreira, F. (2010). The role of inhibition in the production of disfluencies. *Journal Memory & Cognition*, *38*(5), 617–628.

Favreau, M., & Segalowitz, N. (1983). Automatic and controlled processes in the first- and second-language reading of fluent bilinguals. *Memory & Cognition*, *11*(6), 565–574.

Freed, B. F. (2000). Is fluency, like beauty, in the eyes (and ears) of the beholder? In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 243–265). Ann Arbor, MI: University of Michigan.

Gernsbacher, M. A. (1993). Less skilled readers have less efficient suppression mechanisms. *Psychological Science*, *4*(5), 294–298.

Geva, E., & Siegel, L. S. (2000). The role of orthography and cognitive factors in the concurrent development of basic reading skills in bilingual children. *Reading and Writing: An Interdisciplinary Journal*, *12*(1), 1–30.

Gholamain, M., & Geva, E. (1999). Orthographic and cognitive factors in the concurrent development of basic reading skills in English and Persian. *Language Learning*, *49*(2), 183–217.

Gilabert, R. (2007). Effects of manipulating task complexity on self-repairs during L2 oral production. *International Review of Applied Linguistics in Language Teaching (IRAL)*, *45*(3), 215–240.

Gollan, T. H., Slattery, T. J., Goldenberg, D., van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: the frequency-lag hypothesis. *Journal of Experimental Psychology: General*, *140*(2), 186–209.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York: Cambridge University Press.

Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics, 30*(4), 461–473.

Häyrinen, T., Serenius-Sirve, S., & Korkman, M. (1999). *Lukilasse. Lukemisen, kirjoittamisen ja laskemisen seulontatestistö peruskoulun ala-asteen luokille 1–6*. [The Lukilasse graded achievement package for comprehensive school age children.] Helsinki: Psykologien Kustannus Oy.

Jones, M. W., Obregón, M., Kelly, M. L., & Branigan, H. P. (2008). Elucidating the componen processes involved in dyslexic and non-dyslexic reading fluency: An eye-tracking study. *Cognition*, *109*(3), 389–407.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122–149.

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.

Kame'enui, E., & Simmons, D. (2001). Introduction to this special issue: The DNA of reading fluency. *Scientific Studies of Reading*, *5*(3), 203–210.

Kaukonen, S., & Lanu, A.-L. (2005). *Lukemisen osatekijöitä ja ongelmia*. [Components and problems of reading.] Unpublished master's thesis. University of Jyväskylä, Finnish studies.

Kirby, J. R., Parrila, R. K., & Pfeiffer, S. L. (2003). Naming speed and phonological awareness as predictors of reading development. *Journal of Educational Psychology*, *95*(3), 453–464.

Koponen, M., & Riggenbach, H. (2000). Overview: Varying perspectives on fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 5–22) Ann Arbor, MI: University of Michigan.

Kormos, J. (1999). The effect of the speaker variables on the self-correction behaviour of L2 learners. *System*, *27*(2), 207–221.

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

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*, 293–323.

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

Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 25–42). Ann Arbor, MI: University of Michigan.

Lyytinen, H., Ahonen, T., Eklund, K., Guttorm, T., Kulju, P., Laakso, M.-L., Leiwo, M., Leppänen, P., Lyytinen, P., Poikkeus, A.-M., Richardson, U., Torppa, M., & Viholainen, H. (2004). Early development of children at familial risk for dyslexia – Follow-up from birth to school age. *Dyslexia*, *10*(2), 146–178.

MacLeod, C. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*(2), 163–203.

Marian, V., Blumenfeld, H. K., Mizrahi, E., Kania, U., & Cordes, A.-K. (2013). Multilingual Stroop performance: Effects of trilingualism and proficiency on inhibitory control. *International Journal of Multilingualism*, *10*(1), 82–104.

Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: asymmetrical costs of language selection. *Journal of Memory and Language*, *40*(1), 25–40.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100.

Olkkonen, S., Eklund, K., & Leppänen, P. H. T. (submitted). Fluency of lexical access in L1 and L2 in relation to L2 literacy.

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. *EuJAL 2016, 4*, 209–238.

Pollatsek, A., Rayner, K., & Balota, D. A. (1986). Inferences about eye movement control from the perceptual span in reading. *Perception & Psychophysics*, *40*(2), 123–130.

Protopapas, A., Fakou, A., Drakopoulou, S., Skaloumbakas, C., & Mouzaki, A. (2013). What do spelling errors tell us? Classification and analysis of spelling errors of Greek schoolchildren with and without dyslexia. *Reading & Writing: An Interdisciplinary Journal*, *26*(5), 615–646.

Riggenbach, H. (Ed.) (2000). *Perspectives on fluency*. Ann Arbor, MI: University of Michigan.

Salmi, P. (2008). *Nimeäminen ja lukemisvaikeus. Kehityksen ja kuntoutuksen näkökulma*. [Naming and dyslexia: Developmental and training perspectives.] Doctoral dissertation. Jyväskylä Studies in Education, Psychology and Social Research, 345.

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, *14*, 357–385.

Segalowitz, N. (2000). Automaticity and attentional skill in fluent performance. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 200–219). Ann Arbor, MI: University of Michigan Press.

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. NY: Routledge.

Segalowitz, N., & Frenkiel-Fishman, S. (2005). Attention control and ability level in a complex cognitive skill: attention-shifting and second language proficiency. *Memory & Cognition*, *33*(4), 644–653.

Simola, J., Holmqvist, K., & Lindgren, M. (2009). Right visual field advantage in parafoveal processing: Evidence from eye-fixation related potentials. *Brain and Language*, *111*(2), 101–113.

Skehan, P. (2003). Task-based instruction. *Language Teaching*, *36*(1), 1–14.

Snellings, P., Van Gelderen, A., & De Glopper, K. (2002). Lexical retrieval: An aspect of fluent second language production that can be enhanced. *Language Learning*, *52*(4), 723–754.

Stroop, J. R. (1935). Studies of interference in serial verbal reaction. *Journal of Experimental Psychology*, *18*(6), 643–662.

Walczyk, J. J., Marsiglia, C. S., Johns, A. K., & Bryan, K. S. (2004). Children's compensations for poorly automated reading skills. *Discourse Processes, 37*(1), 47–66.

Wolf, M. (1986). Rapid Alternating Stimulus naming in the developmental dyslexias. *Brain and Language*, *27*(2), 360–379.

Received March 18, 2016
Revision received October 7, 2016
Accepted April 6, 2017