



Spelling in Finnish: the case of the double consonant

Eva Marinus¹ · Minna Torppa² · Jarkko Hautala³ · Mikko Aro⁴

Accepted: 9 October 2021 / Published online: 12 November 2021
© The Author(s) 2021

Abstract

Because of its regularity, it is relatively easy to learn to read and spell in Finnish. However, a specific hurdle in spelling acquisition seems to be the doubling of consonant letters. In this study on consonant letter doubling spelling in Finnish children (91 Grade 1 and 191 Grade 2 children), we asked two questions. First, are items with double consonant letters (e.g., “kissa” [‘kis:ɑ] ‘cat’) indeed harder to spell than single consonant items (e.g., “kisa” [‘kisa] ‘contest’)? Second, is consonant doubling harder for stop consonants (e.g., “takki” [‘tɑk:i] ‘coat’) than for continuant consonants (e.g., “kissa” [‘kis:ɑ] ‘cat’)? We found that Finnish children made more errors on items with double consonant letters than on items with single consonant letters and that this effect was larger for stop than for continuant consonant letters. Exploratory analyses showed that these effects were stronger for younger and poorer spellers. Post hoc analyses of the errors made on double consonant items showed that the children predominantly made nonlexical errors (>90%). When they did make a lexical error, these errors typically did not map on the type of errors that would be expected from a corpus analysis of the higher-frequency orthographic neighbors. Overall, lexical influences on spelling of Finnish children seem to be minimal and unpredictable. We discuss two potential reasons why it is more difficult to spell items with double consonant letters than with single consonant letters and suggest how these could be investigated in future research.

Keywords Spelling · Finnish · Spelling development · Consonants · Consonant doubling

Introduction

Results of several cross-linguistic studies have demonstrated the universal underlying processes involved in learning to read and spell in different languages (e.g., Caravolas et al., 2012) and in brain circuitry responsible for reading (Rueckl et al.,

✉ Eva Marinus
eva.marinus@phsz.ch

Extended author information available on the last page of the article

2015). However, recent studies (Georgiou et al., 2020; Zeguers et al., 2018) also show that each language poses specific challenges to its readers and spellers, which makes it challenging to directly compare languages in a cross-linguistic study. Here we propose another approach. Instead of directly comparing languages, we pinpoint a language in which enough instances of a specific phenomenon occur to systematically study the target process. For example, compared to other alphabetic languages, English has a high number of irregular words. As a result, English is an excellent language in which to systematically investigate the regularity effect on learning to read. In contrast, trying to run a systematic study about regularity effects in Italian would be much harder as there are many fewer words with irregular grapheme-phoneme correspondence rules available. As another example, if one wants to investigate the effect of letter position, the language of choice would be Hebrew, which has a high number of anagram words, like “כלבים” (‘dogs’) and “כבלים” (‘cables’). Therefore, Hebrew offers an excellent natural context to study letter-position processing in learning to read (Friedmann et al., 2007). Other examples are the influence of vowel letters on the pronunciation of consonant letters in Russian (Schmalz et al., 2017) and complexity vs. irregularity effects in French and English (Schmalz et al., 2016). By strategically conducting experimental studies in different languages, capitalizing on their specific features and challenges, we can systematically isolate and better understand the different processes which the reading and spelling system in the brain is capable of or how it has adapted itself to the linguistic structures of different languages.

In the current study we use Finnish as our target language to better understand how even very regular mapping of sounds to letters can still lead to specific challenges in spelling. Finnish is one of the most transparent alphabetic orthographies in the world (Seymour et al., 2003). Each letter corresponds to a single phoneme, with only minor exceptions (for a detailed description, see Aro, 2017). As a result, it is relatively easy to learn to read in Finnish, despite its complex morphological system resulting from its agglutinative nature and rich derivational system (Aro, 2017). Letters are introduced in preschool year and formal reading education starts at first grade in the year children turn seven. Most Finnish children can read any word after a few months of reading instruction. From Grade 3 onwards, Finnish children are expected to use reading and spelling for independent learning and learning to spell is also relatively easy (Aro, 2017). The only, and very specific, spelling challenge seems to lie in the doubling of consonant letters (e.g., “takki” [ˈtɑk:i] ‘coat’), resulting in the most common error (i.e., failing to double the consonant letters) in early spelling (Lyytinen et al., 1995). A similar challenge seems to be present in Germanic languages, like German, English, and Dutch. Grade 2, 3 and 4 children typically make more errors in spelling words with double consonant letters (e.g., “bommen” [ˈbɔ.mən], which is Dutch for ‘bombs’) compared to words with single consonant letters (e.g., “bomen” [ˈbo:.mən], which is Dutch for ‘trees’) (Landerl et al., 2005; Snickers-Mommer, 2009).

However, looking in detail at the Finnish spelling system, it is quite interesting that consonant-letter doubling poses a challenge the way it does in Germanic languages. In Germanic languages, the doubling of consonant letters signals a different pronunciation of the preceding vowel (e.g., “hoping” vs. “hopping” in English

or “bomen” [ˈbɔː.mən] ‘trees’ vs. “bommen” [ˈbɔ.mən] ‘bombs’ in Dutch). Or, in other words, consonants are doubled when they are the only letter following a short, stressed vowel in an open syllable. In the Finnish orthography, however, there is never such a conflict between open or closed syllable pronunciation rules (open syllable long, closed syllable short) and the standard phoneme-grapheme correspondence rule (Ziegler et al., 2000). On the contrary, in Finnish words short vowels are always represented by a single vowel letter (e.g., “tuli” [ˈtuli] ‘fire’), whereas long vowels are represented by a double vowel letter (e.g., “tuuli” [ˈtu:li] ‘wind’). Similarly, short consonants are represented by one consonant letter and long consonants by double consonant letters¹. This can be done in two different ways. One possibility is by literally lengthening the pronunciation of the consonant. This which works well for letters with a steady state like “r”, “m” and “n” (e.g., “mummo” [ˈmum:o] ‘grandma’). However, due to their nature, stop consonants (e.g., “k”, “p” and “t”) cannot be lengthened in this way in speech. Instead, for double consonant letter words like “takki” [ˈtæk:i] ‘coat’, Finnish speakers lengthen the stop position (i.e., there is a longer pause) before the explosion of the consonant sound. Suprasegmental features also play a role: the preceding short or long vowel is pronounced for a shorter duration in the case of long consonant sound, and these relative length differences guide categorical perception (Aro, 2017).

As Finnish spellers do not have to deal with additional open and closed syllable rules like their peers learning to spell in Germanic languages, the Finnish language gives a unique opportunity to examine why consonant letter doubling might remain challenging, even without this confound. However, an important first step is to confirm that consonant letter doubling is indeed challenging for Finnish children, and if so, to what degree. The reports in the literature so far are based on reports from practitioners (Aro, 2017). As far as we are aware, there has been only one study that systematically examined spelling of double consonant letter words in Finnish children from Grades 1, 2 and 3. However, this study by Lehtonen et al. (2004) did not directly compare double consonant to single consonant words. Instead, they compared double to mixed consonants in pseudowords (e.g., “maariSSaksi” vs. “luutaSToksi”) and their main aim was to investigate if phoneme length awareness predicted spelling performance in Finnish, which it did. This means, however, that there is still no systematic evidence that the spelling of words with double consonant letters in Finnish is indeed harder than the spelling of words with single consonants. Therefore, the first aim of this study is to run the first systematic study to examine this question.

¹ Strictly taken, the principle of consistently marking the long phonemic quantity of a consonant by doubling the corresponding consonant letter (except for marking the long η-sound with bigraph “ng”) is fully consistent only at the level of morphemes. In some word boundaries and multimorphemic items long consonant quantity can be pronounced, although it is not marked in spelling (due to boundary gemination: e.g., “tule tänne” [tule t:än:ne] ‘come here’). However, such items were not used in our study.

The second aim is to examine if consonant letter doubling is harder for stop consonants (e.g., “takki” [ˈtɑk:i] ‘coat’) than for continuant consonants² (that is, consonant phonemes that can be temporally extended, like “kissa” [ˈkis:ɑ] ‘cat’). Aro (2017) suggests that in the case of stop consonants, the length might be especially difficult to discriminate for beginning spellers, since the long sound is reflected by a longer acoustic silence before the phoneme, before the plosion, instead of a longer duration of the phonemic sound of the consonant (which is impossible for stop consonants). This study will also be the first to systematically investigate this hypothesis. For both our aims we have preregistered our hypotheses and predictions on the open science framework (<https://osf.io/3rgbk/registrations>). In short, we predicted that words with double consonant letters (e.g., “kissa” [ˈkis:ɑ] ‘cat’) will be harder to spell than words with single consonant letters (e.g., “kisa” [ˈkisa] ‘contest’) and that this effect is stronger for words with stop consonant letters (e.g., “takki” [ˈtɑk:i] ‘coat’) than for words with continuant consonant letters (e.g., “kissa” [ˈkis:ɑ] ‘cat’). We will also conduct exploratory analyses into the effect of spelling ability (as judged by the teachers) and examine the different error types that the children make.

Method

Participants

Participants were 91 Grade 1 (4 different classrooms) and 191 Grade 2 (9 different classrooms) children from three different primary schools from the same area in central Finland. In Finland, children start formal schooling early August in the year they turn seven. Reading instruction in Finnish commences in Grade 1. When the data for this study was collected, the children had not yet received any foreign language instruction. We did not register their ages, but since the data were collected in April typically the ages vary between 7 years and 4 months and 9 years and 4 months. Teachers were contacted after school directors confirmed their willingness to participate in the study. Parental consent was passive, that is, the teachers informed the parents about the study and asked them to let them know if their child was not allowed to participate. One of the parents opted out. Although data were collected in one specific area in Finland, we argue that the population is representative for Finnish children in Grades 1 and 2 as both literacy performance (Ustun et al., 2018; Torppa et al., 2016) and teaching instruction (synthetic phonics) are very homogeneous in Finland.

Each classroom received a gift card worth 50 euros as an appreciation for their participation. In addition, the teachers received feedback about the spelling scores of

² Strictly taken, in articulatory terms, {l, m, s} have nothing in common aside from being consonants and in previous studies, ‘continuant’ typically refers to a continuous oral airstream and therefore does not include /m/. However, after considering other terms (‘lingering’, ‘non-stop’, ‘non-plosive’), we decided to stick to continuant, as in that a continuous sound can be produced.

the children in their classroom, the average number of errors in the full sample, and the typical errors made for different types of items.

Materials

Test items were 48 pseudowords, 12 items per condition. We used pseudowords to avoid word frequency effects. The four conditions were: single stop consonant (e.g., “veko” [ˈveko]), single continuant consonant (e.g., “veso” [ˈveso]), double stop consonant (e.g., “vekko” [ˈvek:o]) and double continuant consonant (e.g., “vesso” [ˈves:o]). The items in the four conditions had the same phonemic structure, they only varied in terms of mid-word consonant letters. Moreover, for the same consonant type, we selected consonant pair items (e.g., “veso” [ˈveso] and “vesso” [ˈves:o]). For stop consonants we used “t”, “p” and “k”, for continuant consonants we used “m”, “s” and “l”. The full item list can be found in the [Appendix](#).

The 48 pseudowords were randomly divided between two lists for each classroom except for one rule: we always made sure that single and double consonant members of each pair (e.g., “veso” [ˈveso] and “vesso” [ˈves:o]) were not in the same list. To ascertain that the auditory presentation of the pseudowords was the same for all children we used a recording of the pronunciations. For this recording we used the spoken pronunciations of one of the research assistants, who was in the fourth year of a teacher education program.

Procedure

The study was conducted in April 2018 and took place in the classrooms of the children during regular school hours. The recorded pronunciations of the items were presented to the entire class at once, except for one class, which was split in two. The number of children in a classroom ranged between 12 and 20.

There were two test sessions, which were typically one week apart (range: 2–19 days). In session one we presented the first 24 items in random order. In session two we presented the other 24 items. So, all children spelled all items. Within each session, there was a short break after the first 12 items. During this break the students walked to their teacher’s desk to return the paper and to take the new paper for the last 12 items.

Assessments were always conducted by two research assistants who were trained by the second author. One assistant focused on the administration of the test materials whereas the other assistant focused on the children (e.g., checking if there was enough time for them to write down each item).

The procedure and instructions for the children in the classroom were as follows. First, the research assistants introduced themselves and the study to the children. Next, they gave the following instructions: “Let’s do a dictation exercise. The words you hear are nonsense words that do not mean anything. You hear each word on the tape two times. Please write the words in lines in numerical order. If you miss a word, don’t worry, just try the next one. Please try to do your best and focus on your own work. Do you have any questions?” After questions were clarified, the pronunciations of the

pseudowords were presented to the children via Logitech S150 speakers, attached to a Dell E7250 laptop. During the session children were asked to raise their hands if they could not hear an item well. In addition, there was always a research assistant in the back of the classroom to check if the sound was loud and clear enough. None of the children reported difficulties hearing the items.

Scoring and classification of the responses

The spellings of the children were scored and classified by the second author and two research assistants. For the main preregistered analyses the responses were scored as either fully accurate, incorrect, or missing. We did not distinguish between number and types of error per item. An item was scored ‘incorrect’ from one error onwards. In other words, ‘incorrect’ ranged from the omission of one letter, two wrongly substituted letters or any mix of other types of spelling errors. Based on the number of incorrectly spelled items, we calculated total accuracy percentage per condition. For the target consonant accuracy analyses, it did not matter if there were errors elsewhere in the word; the responses were scored as correct if the marking of the target consonant(s) was/were correct. Based on this score, we calculated the target consonant accuracy per condition.

For the exploratory analyses, we also counted the number of items with other (not target consonant) errors and calculated the percentage of items with other errors per condition. Finally, we classified the error types for the double consonant condition only. Details of this classification are presented in the final section of the exploratory analyses.

Results

The preregistration, raw and processed data and the results of additional analyses can be found on the open science framework (<https://osf.io/3rgbk/registrations>, <https://osf.io/3rgbk/>).

Missing data

There were two types of missing data. First, 21 children attended only one of the two sessions (3.7% of the data). For these children we calculated, per child, total accuracy percentage, target consonant accuracy percentage, and other error accuracy percentage based on 6 instead of 12 items per condition. Note that this is different from what we said in our preregistration (which states that we would remove the data for these participants). However, since there is no reason to assume that one of the lists contained more difficult items than the other, we decided to keep all the data. Importantly, this decision was made before we ran the analyses. Second, altogether for only 26 items, children did not write down any spelling at all (0.20% of the data). We treated these missing values as errors.

Table 1 Overall accuracy percentage for each of the four conditions

Number of consonants	Type	Mean	SD	N
1	Continuant	84.19	14.06	282
1	Stop	84.04	13.56	282
2	Continuant	77.04	21.73	282
2	Stop	78.16	25.98	282

Analyses

We will first present the results for the preregistered research questions and their corresponding planned analyses. After that we describe the results of the exploratory analyses, that is, analyses for which we did not have specific predictions up front. All analyses were conducted in JASP (JASP Team, 2018).

Planned analyses

We conducted a 2×2 Repeated Measures ANOVA with two within-subjects variables. The within-subjects variables were number of consonants (1 or 2) and type of consonant (continuant or stop). We analyzed two outcome measures: total accuracy percentage and accuracy of the target consonant(s).

Total accuracy percentage

As can be seen in Table 1, total accuracy was quite high. There was a main effect for number of consonants $F(1, 281) = 30.47, p < 0.001, \eta_p^2 = 0.098$. Children were less accurate in their spelling of items with double consonant letters than in their spelling of items with a single consonant letter. The main effect for type and the interaction between number of consonant letters and type were not significant. In other words: the accuracy percentage did not differ between items with continuant and stop consonant letters and it was not the case that the effect of number of consonant letters was more pronounced for continuant or for stop consonants.

Target consonant(s) accuracy percentage

As can be seen in Table 2, accuracy was high. Just like for total accuracy percentage, we found a main effect for number of consonant letters, $F(1, 281) = 22.03, p < 0.001, \eta_p^2 = 0.073$. Children were less accurate in their spelling of double consonant letters than in their spelling of single consonant letters. Again, the main effect for type was not significant, $F(1, 281) = 2.11, p = 0.15$. However, we did find a significant interaction effect between number of consonant letters and type, $F(1, 281) = 17.51,$

Table 2 Target consonant accuracy percentage

Number of consonants	Type	Mean	SD	N
1	Continuant	95.57	9.20	282
1	Stop	96.81	9.15	282
2	Continuant	91.14	21.79	282
2	Stop	88.80	25.02	282

$p < 0.001$, $\eta_p^2 = 0.059$. This means that, overall, accuracy for continuant consonant letters did not differ from accuracy for stop consonant letters. However, for stop consonant letters the difference between the single and double consonant conditions was larger than for continuant consonant letters.

Exploratory analyses

We ran two different types of exploratory analyses, that is, we conducted analyses for which we had no specific predictions when we preregistered the study. First, we had mentioned in our preregistration that if we could get standardized test results from the teachers, we would conduct exploratory analyses of the influence of general spelling proficiency on performance on our experimental task. Unfortunately, results of such a task were not available for the children in our participating schools. Instead, we obtained a teacher evaluation of spelling skills of the children, by asking the teachers to rate them on a scale of 1 (low) to 5 (high). We used this variable as a between-subjects variable and added it to the same Repeated Measure ANOVA we conducted for our planned analyses. Next, we ran the same analyses with Grade as a between-subjects variable (see first and second section below) to examine the effect of grade level on the results.

Second, we noticed that children made many other types of errors. For example, despite the instruction that all items were nonsense words, some children still wrote down existing words, indicating potential indirect lexical effects. For example, “musa” [ˈmusɑ] (common spoken language form for “musiikki” ‘music’) instead of “vusa” [ˈvusɑ]. We also observed many instances of ‘other’ errors, that is, errors that did not involve the doubling of the target (middle) consonant letters. Our explorations of these observations, including corpus analyses, will be presented in the third and fourth sections below.

Total-accuracy-percentage: interactions between medial consonant length and type and spelling ratings or grade

Graphic displays of the results of these analyses can be found in Figs. 1 (spelling rating) and 2 (Grade). First, we ran the analyses with spelling rating. To save space we will only report the statistics for effects with spelling rating and Grade here. A document with all outcomes can be found on the OSF (<https://osf.io/3rgbk/>).

Fig. 1 Total accuracy percentage for each of the five teacher spelling ratings (1 = low, 5 = high), showing the 2-way interaction between number of consonants and spelling rating. Error bars (Standard errors) included

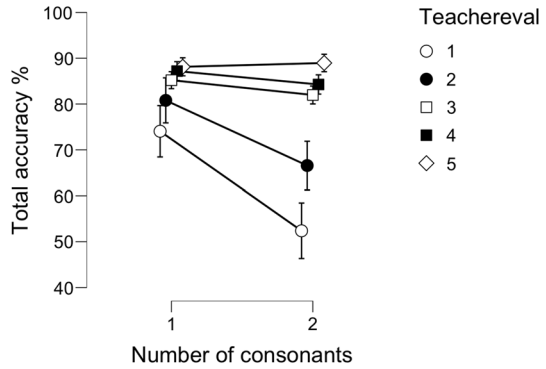
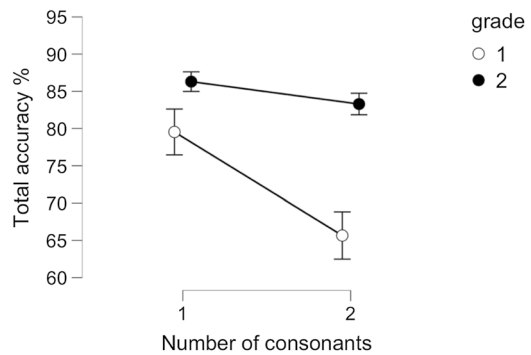
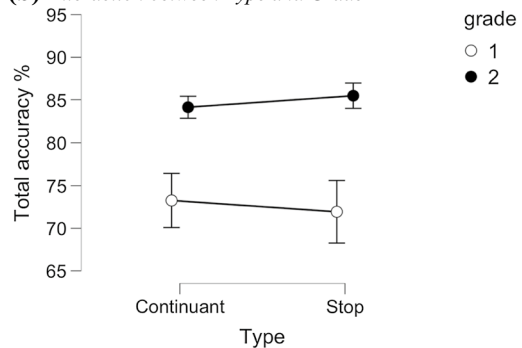


Fig. 2 Total accuracy percentage. Error bars (Standard errors) included. a Interaction between number of consonants and Grade. b. Interaction between Type and Grade

(a) Interaction between number of consonants and Grade



(b) Interaction between Type and Grade



As expected, we found a main effect of number of consonant letters on total spelling accuracy. We also found a significant main effect for teacher evaluation of the spelling skills $F(4, 276) = 24.83, p < 0.001, \eta_p^2 = 0.27$ and a two-way interaction between number of consonants and evaluation of spelling skills, $F(4, 276) = 10.49, p < 0.001, \eta_p^2 = 0.12$. This means that the children who received higher spelling

ratings from their teachers were more accurate in their spelling than the children who received lower ratings. More interestingly, the effect of number of consonant letters was stronger for the lower rated than for the higher rated children. As can be seen in Fig. 1, there was no difference between number of consonant letters for the children that received high ratings. The other effects were not significant ($ps > 0.05$).

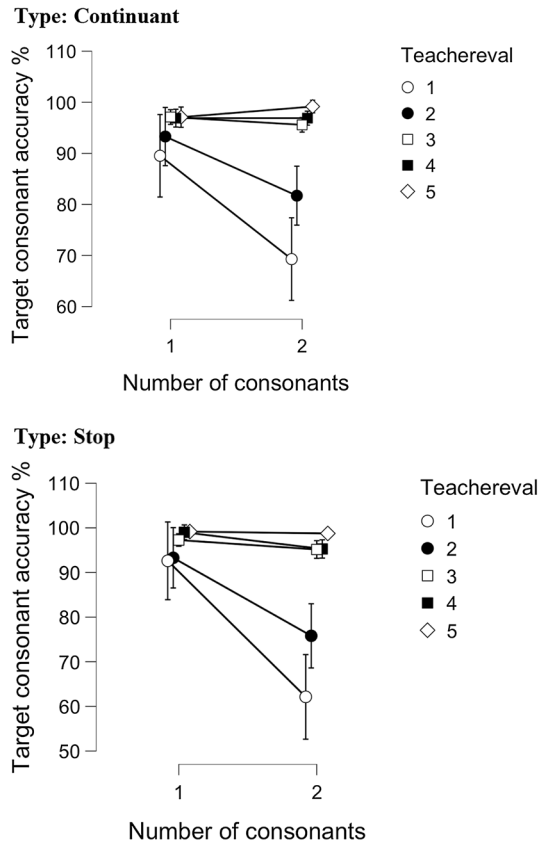
Next, we ran the analyses with Grade. Again, we found a main effect of number of consonants on total spelling accuracy. The main effect for Grade was also significant $F(1, 280) = 44.44, p < 0.001, \eta_p^2 = 0.14$. In addition we found a significant two-way interaction between number of consonants and Grade, $F(1, 280) = 19.74, p < 0.001, \eta_p^2 = 0.057$ and between type and Grade, $F(1, 280) = 5.66, p < 0.05, \eta_p^2 = 0.020$. The children in Grade 2 were more accurate in their spelling than the children in Grade 1. Moreover, the effect of consonant number was stronger for Grade 1 than for the Grade 2 children. The two-way interaction between Type and Grade is harder to interpret, in particular in the absence of a main effect of Type. From Fig. 2b it seems that the difference between Grade 1 and 2 children is bigger for stop than for continuant consonants. The other effects were not significant ($ps > 0.05$).

Target-consonant-accuracy-percentage: interactions between medial consonant length and type and spelling ratings or grade

Graphic displays of the results of these analyses can be found in Figs. 3 (spelling rating) and 4 (Grade). First, we ran the analyses with spelling rating. As expected, we replicated the main effects for number of consonants on target consonant spelling accuracy and the two-way interaction between number of consonants and Type. We also found a main effect for teacher evaluation of the spelling skills $F(4, 276) = 25.24, p < 0.001, \eta_p^2 = 0.27$, a two-way interaction between number of consonants and evaluation of spelling skills, $F(4, 276) = 12.13, p < 0.001, \eta_p^2 = 0.13$, and a three-way interaction between number of consonant letters, type, and spelling rating, $F(4, 276) = 3.37, p < 0.05, \eta_p^2 = 0.043$. As can be seen from Fig. 3, this means that the difference in target consonant letter spelling accuracy between one and two consonant letter items is larger for the poorer than for the better spellers and that this effect is even more pronounced for stop than for continuant consonant letters. The other effects were not significant ($ps > 0.05$).

Next, we ran the analyses with Grade. Again, we found a main effect of number of consonant letters on total spelling accuracy and a two-way interaction between number of consonant letters and type. The main effect for Grade was also significant $F(1, 280) = 34.61, p < 0.001, \eta_p^2 = 0.11$. In addition, we found a two-way interaction between number of consonant letters and Grade, $F(1, 280) = 15.74, p < 0.001, \eta_p^2 = 0.048$ and a significant three-way interaction between number of consonant letters, type and Grade $F(1, 280) = 4.65, p < 0.05, \eta_p^2 = 0.015$. The other effects were not significant ($ps > 0.05$). As can be seen from Fig. 4, this means that the difference in target consonant spelling accuracy between one and two consonant words is larger for the Grade 1 than for the Grade 2 spellers and that this effect is even more pronounced for stop than for continuant consonants. The other effects were not significant ($ps > 0.05$).

Fig. 3 Target consonant accuracy percentage depicting the 3-way interaction between number of consonants, type and teacher-literacy rating. Error bars (Standard errors) included



Other error analyses

Although our research questions are focused on errors in consonant letter doubling, we observed that children made many other errors (e.g., substituting the first letter for another one). If it is the case that children make more other errors in some of our conditions than in the others, this might change the interpretation of the overall accuracy results. We therefore conducted additional exploratory analyses on the other error percentage for each of the four conditions. For this we used again a Repeated Measures ANOVA, with number of consonants and type as within-subjects variables. The descriptives are presented in Table 3.

We found main effects for number of consonants, $F(1, 281) = 12.41, p < 0.001, \eta_p^2 = 0.042$ and type, $F(1, 281) = 50.80, p < 0.001, \eta_p^2 = 0.15$. The interaction effect between number and type was not significant, $F(1, 281) = 2.26, p = 0.13$. This means that children make more other spelling errors for double consonant than for single consonant items and more errors on stop consonant than on continuant consonant items.

Table 3 Other (not related to the middle consonant letter(s)) error percentage per condition

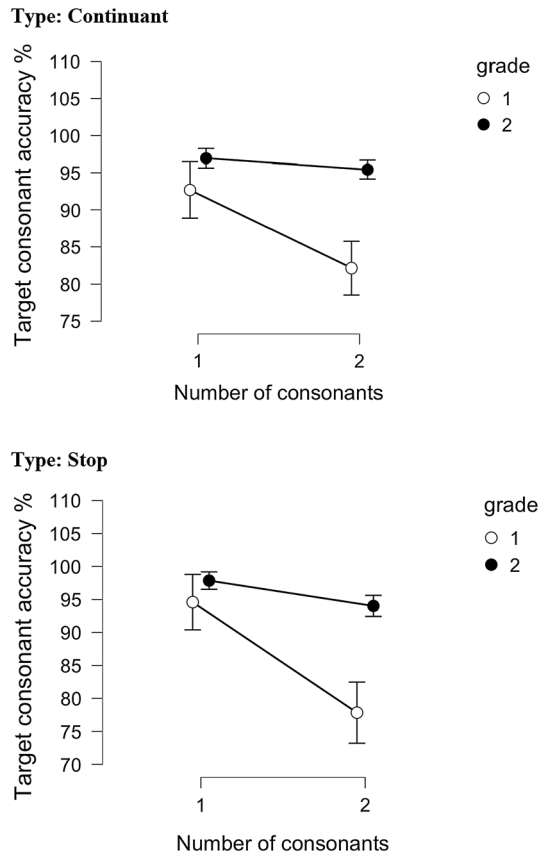
Number of consonants	Type	Mean	SD	N
1	Continuant	5.94	9.74	282
	Stop	8.81	11.65	282
2	Continuant	6.71	9.82	282
	Stop	10.85	11.40	282

Error types on double consonant letter items and corpus analyses

On average, children's error rate on the double consonant letter items was 22.4%. Just under half of this percentage, about 10%, consisted of the expected error of leaving out a consonant letter. The other half, about 11%, mainly consisted of other pseudoword (non-lexical) errors (e.g., "nilo" instead of "milo"), leaving only about 1.4% for lexicalization errors (e.g., "musa" instead of "vusa"). The fact that most errors were pseudowords, either the one consonant-letter variant of the item or another spelling error, seems to suggest that lexical influences do not play a large role in the spelling of Finnish children.

Because we only used pseudowords, one could argue that it makes sense that our results were not so much influenced by lexical effects, such as word frequency. However, lexical effects could still have occurred indirectly, for example via the frequency of the neighbors of our pseudowords. To examine this possibility, we ran a corpus analysis to pinpoint all orthographic substitution neighbors and their frequencies of our pseudowords items (Kotimaisten kielten tutkimuskeskus [Research Institute for the Languages of Finland], 2007). The second and third author (both native speakers of Finnish) checked these neighbors and deleted items that were misspellings or very uncommon. Next, we checked if our experimental conditions differed on number of orthographic neighbors and mean frequency of the orthographic neighbors. This was not the case: there were no significant differences between the experimental conditions for the stimulus items, neither in terms of number of orthographic neighbors, $F(3, 47) = 1.32, p = 0.28$, nor in terms of mean frequency of the orthographic neighbors, $F(3, 39) = 0.39, p = 0.76$. Although this strategy for controlling for neighborhood and frequency effects is often applied, especially in reading research in Germanic languages (e.g., Steacy et al., 2016), it does not seem to be an optimal strategy to control for the spelling of Finnish children. When children made a lexical error, which barely happened, it did not follow from what would be expected from the results of the corpus analyses. For example, based on the corpus data we would expect children to misspell the pseudowords "hykä" [ˈhɣkæ] and "hymä" [ˈhɣmæ] as the real word *hyvä* ('good'). According to the corpus, *hyvä* was the most frequent neighbor word of our items (frequency of 1459 per million) and is

Fig. 4 Target consonant accuracy percentage, showing the 3-way interaction between number of consonants, type and Grade. Error bars (Standard errors) included



obviously a word that children should be familiar with. However, of our entire sample only one child wrote down *hyvä*. Interestingly, the most common lexical error (164 incidences) was misspelling “vusa” [ʼvusa] for *musa* [ʼmusa]. *Musa* only has a frequency of 2 per million in the corpus. Two other words were seen relatively frequently: “vessa” [ʼves:ɑ] ‘toilet’, as a neighbor of the pseudoword “vesso” [ʼves:o] and “mitta” [ʼmit:ɑ] ‘measure’, as a neighbor of the pseudoword “mitto” [ʼmit:o] which respectively appeared 32 and 37 times in the children’s spelling.

Although the additional analyses of higher frequency orthographic neighbors are interesting, it is important to keep in mind that the focus of this research was the spelling accuracy of single and double consonant letters. The data clearly show that this was the main source of spelling errors. When looking for orthographic explanations for errors, it is therefore more relevant to check the general frequency

difference between single (e.g., “kisa” [ˈkisa] ‘contest’) and double (e.g., “kissa” [ˈkis:ɑ] ‘cat’) consonant letter words in Finnish than looking at potential lexical errors originating from neighbor words. If single consonant letter words are more frequent than double consonant letters in general, this could form an alternative explanation for why they are easier for Finnish children to spell. However, a corpus analysis showed that the opposite is true: words with double consonant letters in Finnish are typically of higher frequency than their single consonant letter counterparts³ (Kotimaisten kielten tutkimuskeskus [Research Institute for the Languages of Finland], 2007). In other words, if general frequency would have been important, it would have made the double consonant letter words easier than the single consonant words, which is clearly not the case.

Discussion

This study presents the first systematic investigation of the spelling of double consonants letters in Finnish children. To this end we compared total spelling accuracy of pseudowords with and without double consonant letters and spelling accuracy of the target medial consonant(s) of the same items. For the total accuracy measure, we found that Grade 1 and Grade 2 children made more spelling errors in items with double consonant letters than in items with a single consonant letter. For this general, overall spelling-accuracy measure, we did not find a difference between items with stop and continuant consonant letters. With our more precise target-consonant-accuracy measure we confirmed that the higher error rate on the double consonant letter items was indeed caused by a higher percentage of errors in the doubling of consonant letters as compared to the spelling of single consonant letters. Moreover, with this more fine-grained measure, that is, without the interference of all other possible spelling errors children could make, we also showed it is harder to accurately double stop consonant letters than continuant consonant letters.

In addition to answering these two main preregistered questions, we conducted three further explorations. First, we looked at the effects of Grade and spelling proficiency (as estimated by the teachers). As expected, Grade 2 children were more accurate spellers than Grade 1 children and the children who received higher spelling ratings from their teachers were more accurate than those who received lower spelling ratings. In addition, the difference in overall accuracy between single consonant and double consonant letter items was larger for the Grade 1 than for the Grade 2 and larger for lower-rated children than for higher-rated children. Focusing on the more precise, target-consonant-accuracy measure, we found the same results. Moreover, the effects were stronger for stop than for continuant consonants. Together, these results show that words with double consonant letters are

³ Mean surface word frequencies (in a million word) for 30 identified CVCV-CVCCV pairs (e.g., “kisa”- “kissa”) were 84 for CVCV and 185 for CVCCV counterparts, derived from a newspaper corpus (Research Institute for the Languages of Finland, 2007).

challenging for younger and poorer spellers, in particular if they require the doubling of stop consonants.

Our second exploration focussed on our observation that children made many other errors (e.g., substituting the first letter for another one, like “vusa” [ˈvusa]—*musa* [ˈmusa]). We wanted to further explore this observation, because higher numbers of other, unexpected errors might have influenced the interpretation of the overall accuracy outcomes. For example, high numbers of other errors in the continuant condition could have cancelled out effects of consonant type. However, this did not seem to be the case as we found that children happened to make more other errors on items with stop consonant letters than on items with continuant consonant letters. Therefore, if anything, this should have enhanced the chance to find an effect for consonant type, which was not the case.

However, for consonant letter *number* (single vs. double) we did find an effect in the same direction as we did for our overall accuracy analyses: children also made more other spelling errors for items with double consonant letters than for items with a single consonant letter. One might therefore argue that the main effect of number of consonant letters (i.e., words with double consonant letters are more difficult than those with a single consonant letter) for overall accuracy was mainly caused by other errors. However, we think that this is highly unlikely, as our more fine-grained analyses specifically demonstrated children made more errors with double consonant letters. We therefore speculate that the higher number of other errors for items with double consonant letters reflects that these items produce additional cognitive load for the children, leading them to get confused and make other spelling errors in the process.

Our third and final exploration consisted of error and corpus analyses of the double consonant items with the aim to find out if lexical or frequency factors cause Finnish children to experience difficulties with the spelling of the double consonant letter items in this study. However, this did not seem to be the case. Children barely made lexical errors (writing existing words instead of the dictated pseudoword) and the ones that occurred were not what would be expected based on existing neighbor words. In addition, a corpus analysis showed that words with double consonant letters in Finnish are typically of higher frequency than their single consonant letter counterparts. If lexical factors would have had a key influence on the spelling of Finnish children, they should have found the double consonant letter words easier than the single consonant words.

Then what could be causing the spelling difficulty with double consonants in the medial position? And what could explain that doubling is even harder for stop consonants? Could it be that younger and certain children find it hard to perceive differences in phoneme length and that this causes them to choose the incorrect number of consonant letters? Such an explanation would fit the finding of Lehtonen et al. (2004) that phoneme length awareness in Finnish children is more strongly related to spelling performance than phoneme identity, or phoneme quality, awareness. However, if it were indeed the case that Finnish beginning or struggling spellers are having difficulties *hearing* the difference between the phonemes corresponding to single and double consonant letters, one would expect them to randomly double or not double the consonants. Our results, however, show that this was not the

case: children were more likely to leave out a consonant than to add one. Moreover, speech perception research has convincingly ruled out that categorical perception difficulties play a role in reading and spelling problems (e.g., Zhang et al., 2010). Lastly, one would expect that problems in categorical auditory perception of length would then also lead to widespread problems in spoken language in a language environment like Finnish in which the perception of phonemic length is so important, which is clearly not the case.

A better explanation for the difficulty with double consonant letters seems to be that, although beginning and poor spellers are perfectly able to perceive the auditory differences between single and double consonant letters, they are still skipping this step in the spelling processes. In other words, they are focusing on phoneme quality and overlook its quantity. In a similar vein, Lehtonen et al. (2004) argued that the spelling of long phonemes involves overcoming two hurdles. First, to overcome the assumption that one sound belongs to one letter, that is, the default strategy of using only one consonant letter in the spelling of both short and long consonant phonemes. Such a default strategy could be explained by the fact that overall, when position in the word is not considered, consonant letters occur more often in single than in double formation. Also, the fact that long phonemic quantity of consonants is not always marked in spelling of multimorphemic items, or at word boundaries (see Footnote 1) might also bias young spellers to rely on single letter marking. The second hurdle is to determine and monitor phoneme quantity, which depends on the phonemic context. Experimental training studies are needed to tease apart the importance and relevance of these two different potential hurdles.

To examine if children are using a one-consonant letter default strategy, we could conduct a training experiment that exposes the children to many words with double consonant letters, asking them to mark the double letters and copy the consonant clusters and words. If we find improvement in their spelling after such a treatment, this could be considered evidence for the importance of overcoming a default strategy of marking the phoneme with one letter, regardless of its quantity.

To examine if they accurately monitor consonant phoneme length, telling children to listen better whether they hear a short or long sound might be helpful for continuant consonant letters (e.g., “m”), but not for stop consonants. This is because the explosion of stop consonants is identical, irrespective of phonemic quantity. Therefore, in addition to an ‘encourage active listening’ condition, we would argue for an experiment that also includes a ‘focus on kinetic aspects of articulation’ condition. In this condition, children would not only be instructed to repeat the word they are spelling, but also to pay attention to whether their tongue or lips hold the explosion of the sound for a short or a long period of time. If such instructions improve their spelling, we would have found evidence for the importance of monitoring phoneme

quantity in the spelling of double consonant phonemes. Moreover, we would predict different outcomes for the different training conditions. For the doubling of stop consonant letters, we would only predict better outcomes for the children who attended the ‘focus on kinetic aspects of articulation’ condition. For the doubling of continuant consonant letters, we would predict improvement for in both training conditions.

Limitations

In this study, we were able to answer the preregistered research questions reasonably straightforwardly. However, there is always room for improvement of the design and additional questions came up during the study. For example, we administered the spelling task in group sessions. For future studies, we would recommend individual sessions, in which the participants are asked to repeat the items, to avoid potential perception error confounds. For the Finnish population, we can confidently say that there is not much variation between schools in teaching methods and literacy performance of the students, allowing us to generalize results across schools and Finland. In addition, children are not segregated in low-performance and high-performance classrooms within schools. However, if similar studies were conducted in other countries, researchers could consider controlling for school or even classroom and to carefully select the schools across the country when aiming to generalize beyond the specific area. In our study we used pseudowords because, in contrast to words, they are unfamiliar to all participants and because it enabled us to find matching items for all the conditions. This would not have been possible to the same degree with existing words. However, it needs to be acknowledged that we are generalizing our results with pseudowords to the observation that children struggle with double consonant *words*. Finally, for our exploratory analyses, we used teacher estimates as a measure for literacy performance. Future studies could consider obtaining direct literacy measures (e.g., standardized measures of reading and spelling performance) from the participants.

Conclusion

Our study with beginning spellers of Finnish showed that doubling consonant letters, and in particular stop consonants, forms a specific challenge in learning to spell. The effects were strongest for the younger and poorer spellers. Most primary school teachers in Finland are probably at least partly aware of the double-consonant letter challenge in spelling. However, the results of our systematic study underlines that they need to be specifically vigilant when the spelling of such words is introduced. The examination of error types showed that children predominantly make nonlexical errors and that lexical influences on their spelling seem to be minimal and unpredictable. We argue that the consonant-letter-doubling difficulty can probably be explained by a combination of adopting a default strategy of only using single consonant letter spelling and poor monitoring of phoneme length (or preceding pause in

the case of stop consonants) during the spelling process. We have made suggestions of how experimental training studies could examine these two explanations.

Appendix

Stop consonants		Continuant consonants	
Single	Double	Single	Double
hykä ['hykæ]	hykkä ['hyk:æ]	hymä ['hymæ]	hymmä ['hym:æ]
veko ['veko]	vekko ['vek:o]	veso ['veso]	vesso ['ves:o]
kite ['kite]	kitte ['kit:e]	kime ['kime]	kimme ['kim:e]
rate ['rate]	ratte ['rat:e]	räle ['ræl:e]	ralle ['ræl:e]
vupa ['vupa]	vuppa ['vup:ɑ]	vusa ['vusɑ]	vussa ['vus:ɑ]
jöpä ['jöpæ]	jöppä ['jöp:æ]	jölä ['jølæ]	jöllä ['jøl:æ]
tike ['tike]	tikke ['tik:e]	time ['time]	timme ['tim:e]
soke ['soke]	sokke ['sok:e]	sole ['sole]	solle ['sol:e]
mito ['mito]	mitto ['mit:o]	milo ['milo]	millo ['mil:o]
nöty ['nöty]	nötty ['nöt:y]	nösy ['nøsy]	nössy ['nøs:y]
nopu ['nopu]	noppu ['nop:u]	nomu ['nomu]	nommu ['nom:u]
mäpö ['mäpø]	mäppö ['mäp:ø]	mäsö ['mæsø]	mässö ['mæs:ø]

Funding and Acknowledgements E.M. was supported by the Australian Research Council Centre of Excellence for Cognitions and its Disorders (CE110001021), www.ccd.edu.au. M.T. and J.H. were supported by the Academy of Finland (276239, 317030). Open Access funding was provided by the institution of E.M., the Pädagogische Hochschule Schwyz. We would like to thank the participating schools, teachers and children and Essi Alaluusua, Mirka Joro, Elisa Korhonen, Venla Pousi, and Soile Runtti for their help with the data collection in the schools.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References


- Aro, M. (2017). Learning to read Finnish. In L. Verhoeven & C. Perfetti (Eds.), *Learning to read across languages and writing systems* (pp. 416–436). <https://doi.org/10.1017/9781316155752.017>
- Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavsky, M., Onochie-Quintanilla, E., Salas, N., Schöffelová, M., Defior, S., Mikulajová, M., & Seidlová-Málková Hulme, C. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science*, 23(6), 678–686.

- Friedmann, N., & Rahamim, E. (2007). Developmental letter position dyslexia. *Journal of Neuropsychology*, 1(Pt 2), 201–236. <https://doi.org/10.1348/174866407X204227>
- Georgiou, G. K., Torppa, M., Landerl, K., Desrochers, A., Manolitsis, G., Jong, P. F., & d., & Parrila, R. (2020). Reading and spelling development across languages varying in orthographic consistency: Do their paths cross? *Child Development*, 91(2), e266–e279. <https://doi.org/10.1111/cdev.13218>
- Hautala, J., Aro, M., Eklund, K., Lerkkanen, M.-K., & Lyytinen, H. (2013). The role of letters and syllables in typical and dysfluent reading in a transparent orthography. *Reading and Writing*, 26(6), 845–864. <https://doi.org/10.1007/s11145-012-9394-3>
- Hautala, J., Hyönä, J., Aro, M., & Lyytinen, H. (2011). Sublexical effects on eye movements during repeated reading of words and pseudowords in Finnish. *Psychology of Language and Communication*, 15(2), 129–149. <https://doi.org/10.2478/v10057-011-0009-x>
- JASP Team. (2018). JASP [Computer software] (Version 0.9).
- Kotimaisten kielten tutkimuskeskus [Research Institute for the Languages of Finland]. (2007). *Kirjoitetun suomen kielen sanojen taajuuksia. [Frequencies of written Finnish words]*. <http://kaino.kotus.fi/sanat/taajuuksilista/parole.php>
- Landerl, K., & Reitsma, P. (2005). Phonological and morphological consistency in the acquisition of vowel duration spelling in Dutch and German. *Journal of Experimental Child Psychology*, 92(4), 322–344. <https://doi.org/10.1016/j.jecp.2005.04.005>
- Lehtonen, A., & Bryant, P. (2004). Length awareness predicts spelling skills in Finnish. *Reading and Writing*, 17(9), 875–890. <https://doi.org/10.1007/s11145-004-2802-6>
- Lyytinen, H., Leinonen, S., Nikula, M., Aro, M., & Leiwo, M. (1995). In search of the core features of dyslexia: Observations concerning dyslexia in the highly orthographically regular Finnish language. In V. W. Berninger (Ed.), *The varieties of orthographic knowledge* (pp. 177–204). Springer.
- Rueckl, J. G., Paz-Alonso, P. M., Molfese, P. J., Kuo, W.-J., Bick, A., Frost, S. J., Hancock, R., Wu, D. H., Einar Mencl, W., Duñabeitia, J. A., Lee, J.-R., Oliver, M., Zevin, J. D., Hoefft, F., Carreiras, M., Tzeng, O. J. L., Pugh, K. R., & Frost, R. (2015). Universal brain signature of proficient reading: Evidence from four contrasting languages. *Proceedings of the National Academy of Sciences*, 112(50), 15510. <https://doi.org/10.1073/pnas.1509321112>
- Schmalz, X., Beyersmann, E., Cavalli, E., & Marinus, E. (2016). Unpredictability and complexity of print-to-speech correspondences increase reliance on lexical processes: More evidence for the orthographic depth hypothesis. *Journal of Cognitive Psychology*, 28(6), 658–672. <https://doi.org/10.1080/20445911.2016.1182172>
- Schmalz, X., Porshnev, A., & Marinus, E. (2017). Two distinct parsing stages in nonword reading aloud: Evidence from Russian. *Quarterly Journal of Experimental Psychology*, 70(12), 2548–2559. <https://doi.org/10.1080/17470218.2016.1247895>
- Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94(2), 143–174. <https://doi.org/10.1348/000712603321661859>
- Snijkers-Mommer, S. (2009). *Liever lessen lezen en spelen spellen: Lezen en spellen van bisyllabische woorden met open en gesloten lettergrepen [‘Better’ read better and “rather” spell rather: Reading and spelling of bisyllabic words with open and closed syllables]* (Master Thesis). University of Amsterdam, Amsterdam.
- Steady, L. M., Kearns, D. M., Gilbert, J. K., Compton, D. L., Cho, E., Lindstrom, E. R., & Collins, A. A. (2016). Exploring individual differences in irregular word recognition among children with early-emerging and late-emerging word reading difficulty. *Journal of Educational Psychology*. <https://doi.org/10.1037/edu0000113>
- Torppa, M., Georgiou, G. K., Lerkkanen, M. K., Niemi, P., Poikkeus, A. M., & Nurmi, J. E. (2016). Examining the simple view of reading in a transparent orthography: A longitudinal study from kindergarten to grade 3. *Merrill-Palmer Quarterly*, 62(2), 179–206. <https://doi.org/10.1177/0741932518762055>
- Ustun, U., & Eryilmaz, A. (2018). Analysis of finnish education system to question the reasons behind finnish success in PISA. *Studies in Educational Research and Development*, 2(2), 93–114.
- Zegers, M. H. T., van den Boer, M., Snellings, P., & de Jong, P. F. (2018). Universal and language-specific predictors of early word reading in a foreign language: An analysis of the skills that underlie reading acquisition in three different orthographies. *Developmental Psychology*, 54(12), 2274–2290. <https://doi.org/10.1037/dev0000606>

- Zhang, J., & McBride-Chang, C. (2010). Auditory sensitivity, speech perception, and reading development and impairment. *Educational Psychology Review*, 22(3), 323–338. <https://doi.org/10.1007/s10648-010-9137-4>
- Ziegler, J. C., Perry, C., & Coltheart, M. (2000). The DRC model of visual word recognition and reading aloud: An extension to German. *European Journal of Cognitive Psychology*, 12(3), 413–430. <https://doi.org/10.1080/09541440050114570>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Eva Marinus¹  · Minna Torppa² · Jarkko Hautala³ · Mikko Aro⁴

Minna Torppa
minna.p.torppa@jyu.fi

Jarkko Hautala
jahaut@nmi.fi

Mikko Aro
mikko.t.aro@jyu.fi

- ¹ Pädagogische Hochschule Schwyz, Goldau, Switzerland
- ² Department of Teacher Education, University of Jyväskylä, Jyväskylä, Finland
- ³ Niilo Mäki Institute, Jyväskylä, Finland
- ⁴ Department of Education, University of Jyväskylä, Jyväskylä, Finland