

Running head: Sources of information in learning to spell

The Sources of Information Children Use in Learning to Spell:
The Case of Finnish Gemimates

Annukka Lehtonen

Department of Experimental Psychology

University of Oxford, U.K. *

* Now at Washington University in St. Louis, U.S.A.

Abstract

Two experiments were designed to investigate the ways in which children use orthographic, phonological and morphological information in spelling double consonants (geminate) in Finnish. In the first experiment, children had to choose out of two pseudo-word spellings the one that looked more like a real word on the basis of orthographic or phonological information. In the second experiment, children spelled real words containing target consonant clusters in either word stems or inflections. The results showed that even children just starting school were able to use orthographic information to their advantage in spelling, while phonological aspects of spelling rules were only acquired later on. During the first school year also the use of morphological information began to emerge. Thus, children seem to use multiple kinds of information in spelling from very early on.

1 Introduction

Traditionally, the study of literacy development has been dominated by research concentrating on the English orthography. However, languages differ in their structure and the characteristics of their orthography, and English with its deep morphophonological orthography and relatively simple morphological structure is by no means the norm. This is increasingly acknowledged in the field, as the past decade has seen a remarkable increase in the number of studies investigating literacy acquisition in languages other than English, including studies concerning bilingualism and cross-linguistic comparisons in a variety of languages. Because languages differ from each other in multiple ways, they provide different kinds of opportunities for investigating the factors affecting literacy development. Finnish is an interesting language for this kind of research for several reasons. It has an almost perfectly transparent orthography, phoneme length in addition to phoneme quality distinguishes between meanings, and the morphological structure of the language is very complex.

1.1 Characteristics of Finnish

1.1.1 Orthography

The most remarkable characteristic of the Finnish orthography is that it is very transparent, because the relationships between graphemes and phonemes are regular and consistent. This is true for both spelling-to-sound and sound-to-spelling relationships, so that each letter denotes only one sound and each sound can only be spelled with one letter. This makes Finnish symmetrically transparent, unlike for example French and German, which are more transparent in their spelling-to-sound relationships than in their sound-to-spelling relationships and thus easier to read than to spell. The only exception to the one-to-one phoneme-grapheme mappings in Finnish is the nasal consonant phoneme /ŋ/, which is represented by the digraph *ng*. Finnish uses the Roman alphabet to

represent its 13 consonants and 8 vowels. It is worth noting that Finnish has fewer phonemes than for example English, as there are only 21 phonemes compared to the 44 that are used in English. Therefore, not all the letters of the Roman alphabet are used in Finnish, e.g. *c* and *w* are never used to represent Finnish phonemes. Consequently, Finnish gives young readers and spellers two kinds of advantages in acquiring the alphabetic principle. Firstly, the number of different sounds that need to be distinguished and represented is smaller than in languages that are richer in phonemes. Secondly, the way in which phonemes are mapped onto letters is much more straightforward than in most other orthographies. A further constraining factor is the small number of legal syllable structures allowed in Finnish. Finnish only allows 14 basic syllable types, and therefore syllables provide quite stringent phonotactic constraints for Finnish words and possibly give guidelines for beginning and more advanced spellers.

1.1.2 Phoneme length in Finnish

While many languages, such as English, only code for phoneme quality, in Finnish it is also necessary to pay attention to phoneme quantity, i.e. the duration of phonemes in terms of time. There is a difference between short phonemes, which are spelled with a single letter, and long phonemes, which are spelled with two identical letters. The distinction in Finnish is like the difference in duration in pronouncing the letter *a* in the English words *bat* and *bad* (Treiman, 1993), rather than the qualitative difference in pronouncing the same letter in the words *hat* and *hate*. The difference between *bat* and *bad* does not distinguish between meaning in English, but does so in Finnish. Both consonants and vowels can appear as long in Finnish. Long consonants are referred to as geminates. Examples of how the length of phonemes distinguishes between meanings are provided in Table 1.

INSERT TABLE 1 ABOUT HERE

While determining phoneme quality is relatively straightforward in Finnish, defining the length of a phoneme is more complicated, since length is not an absolute concept. Rather, it depends on the total duration of the word in question and also on the other phonemes in the word (Lehtonen, 1970). However, there is no evidence as of yet that normally developing Finnish school-age children have problems distinguishing between short and long phonemes in speech, while children often make errors in representing long phonemes with two letters at the early phases of literacy acquisition. Problems in representation of length also provide one potential marker of reading and spelling problems in Finnish, both for children and adults (Lyytinen, Leinonen, Nikula, Aro & Leiwo, 1995). This might be partly due to the fact that the spelling of long phonemes deviates from the “one sound, one letter” –rule that otherwise holds well in the Finnish orthography because of its transparent nature. Therefore, we can say that representing phoneme length is an inconsistent aspect of this otherwise very regular orthography.

1.1.3 Morphology

Finnish is a very complex language with respect to morphological structure. Possession, plurals, prepositions and some particles are all expressed by inflections added to the ends of words, so Finnish is a good example of an agglutinative language (see Table 2). In practice, this means that words are long and can take on a vast number of different forms. Each noun can have over 2000 different forms, and each verb as many as 10 000. The morphology is particularly complex as words often have several different stems, which are used depending on the type of inflection that is attached to the end of the word in different occasions. Table 3 shows an example of this.

INSERT TABLES 2 AND 3 ABOUT HERE

This kind of morphological structure sets certain requirements for how it can be processed. Niemi & Laine (1995) point out that there simply is not enough time to go through all the possible forms of words if they were all represented as separate entries in the lexicon. Indeed, several studies

looking at reading in Italian (Caramazza, Laudana & Romani, 1988), which is another highly inflected and agglutinative language, and Finnish (Niemi, Laine & Tuominen, 1994) suggest that word stems and inflections are represented separately in the lexicon of Italian and Finnish readers. This work is still at a preliminary stage, however, and we do not know how the representations develop or how they may affect spelling or be affected by spelling.

Most of the aforementioned research in Finnish has investigated case inflections. Finnish has 14 different case inflections, and they are used instead of word order to express the role of a word in a sentence. They are very frequent in the language and children acquire them relatively early. Thus, if morphological information were to influence children's literacy acquisition, case inflections would probably be involved and this is why we decided to concentrate on studying the possible role of case inflections in spelling.

1.2 The types of information that children use in spelling

The development of children's spelling skills is generally considered an increasingly sophisticated process of understanding how the sounds of words can be represented by letters or letter groups. The specifics of this developmental progression have been outlined in several spelling models, for example those of Frith (1985) and Ehri (1992). The different models agree on many of the main characteristics of spelling development. Initially, children do not attempt to represent the sounds in words, but their "spellings" are letters and numbers jumbled together with no clear correspondence to the words that they are supposed to represent. Children's first attempts at representing the sound structure of words are incomplete, as they represent some of the correct sounds, but not all of them. Both letter-name knowledge and children's phonological knowledge affect the outcome of these early spelling attempts. The next step is considered to be the "phonetic" stage of spelling, where children's spellings represent the complete sound structure of the words. However, they still fail to consider many of the conventional constraints of the orthography of the

language (orthographic rules), such as for example legal letter position and the types of letters that can appear adjacent to each other. Children also overlook many morphologically regular spelling sequences. The ability to use orthographic and morphological information in spelling is considered to be a more sophisticated approach on the way to proficiency than the children in the semiphonetic or phonetic stage are capable of. A question that the spelling models disagree on and that has not been resolved yet is whether children use qualitatively different spelling strategies at different times. If this were the case, we would expect to find that children use different kinds of information in spelling at different points in time. In contrast, if we were to learn that children can use different types of knowledge to their advantage in spelling from early on, this would be evidence against the strict stage model approach.

One way to investigate how children's ability to use different kinds of knowledge affects their developing spelling skills is to look at how a specific aspect of spelling develops over time. Cassar & Treiman's (1997) study of children's knowledge of the use of letter doublets offers a good example of this approach. They used an orthographic constraints test, in which children are shown non-word pairs and asked to choose the member of the pair that looks more like a real word. Cassar & Treiman (1997) compared legal and illegal doublets in final (legal) and initial (illegal) positions of the word. They also ran an experiment where the non-words were read out loud, so that it was possible to test children's knowledge of the phonological rule associated with letter doublets, i.e. that a consonant letter doublet usually follows a short vowel (Carney, 1994). The results showed that first-graders chose non-words with final doublets and legal doublets more often than non-words with initial or illegal doublets. Kindergartners were similarly sensitive to the legal position of doublets, while still at chance when having to choose between legal and illegal doublets. This suggests that already very young children know something about the orthographic rules that govern letter doublet use. However, it was not until sixth grade and above that children had begun to master the phonological rule of doublet spelling.

These results illustrate two important issues. Firstly, even kindergartners can make certain judgements on the basis of orthography, although their spelling is still on the semiphonetic or phonetic level. Although these children have not been learning to spell for very long yet, they have been exposed to print in their environment, and have had the opportunity to pick up information about the conventions of their orthography implicitly. Secondly, we can see that information of the function of spelling sequences is acquired gradually, depending on the nature of the information involved. While already kindergartners seemed to know about the allowed position of doublets, it takes an additional year to work out that not all letters are allowed to double. More sophisticated aspects of spelling, like the phonological relationship between the medial doublet and the preceding vowel, are not mastered until considerably later.

Pacton, Perruchet, Fayol & Cleeremans (2001) investigated French children's performance in the orthographic constraints test. They controlled the frequency of letter doublets by using consonants that differ in their frequency of doubling. This is possible since there are consonants (e.g. c, d, v) that are frequent as single letters, but double rarely, while other consonants (e.g. l, m, s) are frequent both as single and double letters. Pacton et al. (2001) also investigated children's learning of the rule about the doublet position by using doublets of letters that are never doubled. Thus, children's responses should not be biased by the number of times that they had seen these doublets in either initial or final position. The results showed that even when the frequency of the single consonants making up the doublets was controlled, children preferred more frequent doublets, and thus showed sensitivity to the frequency of doublets per se already in their first school year. This sensitivity increased from grade 2 to grade 3. Moreover, children were more likely to choose medial doublets (legal in French) than initial or final doublets (illegal in French) even when all the doublets were made up of consonants that cannot be doubled in French and children could not have responded on the basis of how often they had seen the doublets. These results corroborate those of Cassar & Treiman (1997) and suggest that children in the semiphonetic and phonetic stages

of spelling development already possess some orthographic knowledge that they can bring into their spelling.

Before we draw any firm conclusions about children's concept of doublets in spelling, it is important to note that doublets have different functions in different orthographies. In English, doublets usually follow a short vowel, and in principle this should also help children to know that doublets are not allowed in the initial position of a word. In French doublets do not have a phonological function at all, so their role is exclusively orthographic. Yet another type of function appears in Finnish, where doublets represent the length of phonemes. Italian doublets have a similar function, but the phoneme-grapheme relationships in Italian are somewhat more irregular than those of Finnish. These cross-linguistic differences offer an interesting opportunity to investigate the interaction between orthography and phonology. French children have only orthographic information to guide them in the use of doublets, which allows us to look at the development of orthographic information alone. English children have to learn quite a complex rule about doublets, involving both phonology and orthography, and indeed they only begin to follow it at the age of 12 years and above. However, already beginning spellers seem to know something about the consistent rule that specifies the legal position of doublets, i.e. that doublets are not allowed in the initial position of a word.

In Finnish, however, the relationship between consonant doublets (geminate) in orthography and phonology is not consistent with respect to the geminate position within a word. A long consonant in the middle of the word is always marked by a geminate. However, if we look at the initial position of the word, there are word forms (e.g. singular second person imperative: "*Come here*", "*Eat this*") that will make the first consonant of the following word lengthen in pronunciation. This lengthening in the initial position is not represented in spelling, however; initial geminates are never allowed in the Finnish orthography. The reason for this phenomenon is historical. These word forms used to end in a consonant, and although this is no longer marked in

spelling, it is realised in pronunciation (Laaksonen & Lieko, 1998). This means that children cannot simply follow the phonological rule in spelling. Instead, they have to take into account the orthographic constraint that geminates are illegal at word beginnings.

We have investigated how children's knowledge of the orthographic and phonological aspects of the geminate spelling rule affects their performance in an orthographic constraints test, and whether the two types of knowledge interact. We compared children's knowledge of the allowed position of geminates, since the orthographic cue of the legal position of geminates is very straightforward in Finnish, but not explicitly taught.

If Finnish children are sensitive to orthographic constraints like English-speaking children are, we would expect them to be less likely to accept geminates in the initial than in the medial position of words. It is not obvious, however, that we would see this pattern of responding in the first-year children, since at the time of testing they had not yet been exposed to any formal literacy instruction. Therefore, any knowledge they may demonstrate is more likely to be due to informal exposure to print. Additionally, we aimed to investigate children's mastery of the phonological aspect of the geminate spelling rule, i.e. that geminates are used to mark long consonants in speech. We did not expect first-year children to make the distinction between long and short consonants yet, as this usually emerges in spelling only some months after the start of formal literacy instruction. However, the better at spelling the children become, the more attention they should pay to the length of the phonemes that they hear.

The older children's performance is of interest, because their errors could show us whether phonological and orthographic information interact in the course of spelling development. These children have had plenty of experience of reading and spelling already. Thus, they are likely to know that initial doublets are not allowed, but also that long phonemes should be represented by geminates. Therefore, they actually have two types of conflicting information that might lead them to accept disproportionately many word-initial geminates. Third-year children should be less likely

than second-year children to show the effect of the conflict between the types of information, since they are much more competent in spelling in general and their knowledge of the spelling system is far more extensive.

As long consonants are only pronounced at the beginning of the word if they are preceded by certain types of words, it was necessary to use a modified version of the orthographic constraints test, i.e. short sentences where the target pseudo-word was always the second word. We tested children in years 1-3 at the beginning of the school year, when the first-year children have not received any formal instruction in reading or spelling yet. Therefore, it was necessary to use a test where the requirements for reading and spelling were minimized. We solved this problem by reading out loud each sentence and the pseudo-word in it, then asking children to circle, out of two choices, the pseudo-word spelling that looked more like a real word. This way, children did not need to know how to read that well, and as they only had to circle one of the options, they did not have to write anything either. This ensured that we managed to gather data from all first-year children, regardless of the fact that their reading and spelling skills were very limited at this stage.

To investigate the development of and interaction between children's orthographic and phonological knowledge, the test had four conditions. The test sentences included words with a long/short phoneme pronounced at the beginning of some target words and a long/short phoneme in the middle of other target words. The phonemes in the initial position always had to be spelled with one letter, while a long phoneme in the middle of the word was spelled with a letter doublet and a short phoneme with a single letter. Thus, orthography and phonology are in conflict in the condition where a long phoneme is pronounced at the beginning of the word.

2 Study 1

2.1 Method

2.1.1 Participants

We tested 121 children in years 1-3 in a primary school in Espoo, Finland. The mean ages of the children were 7;2 years, 8;2 years and 9;2 years. The reason for the high mean ages is that in Finland children go to school in the autumn of the year when they turn seven. The children were tested in August, ten days after the schools had started in the new academic year, so the first year children had not received any formal literacy training yet.

2.1.2 The task and the procedure

We saw the children as a group in a single experimental session and administered the task to the whole class at once, so about 25 children were doing the task simultaneously.

2.1.2.1 The task

The orthographic constraints test we used consisted of 24 sentences. Each of these was three words long and contained a pseudo-word as the second word. The pseudo-words varied according to two variables: the length of the target phoneme and the position of the target phoneme. In half of the words, the target phoneme was long in pronunciation, in half it was short. Half of the words had the target phoneme in the initial position of the word, while in half of the words the target phoneme was in the middle of the word. The most interesting condition was the one where a long phoneme appeared in the initial position and was spelled with a single letter, since here phonology and orthography contradicted each other. There was no equivalent incongruence in the target phonemes that were in the medial position of the word. Table 4 illustrates the design of the task. The target phoneme is printed in bold, and there is only one condition where a geminate is appropriate. Each

experimental word was five phonemes long and with each item the children had a choice between a five-letter word and a six-letter word. The reliability of the task was .78 (Cronbach's alpha).

INSERT TABLE 4 ABOUT HERE

We had two versions of the task with the items in different randomised orders to avoid order effects. In half of the items, the correct choice was on the right hand side of the response sheet, in half of the items it was the left-hand side word. To make sure that children were paying attention to the right words of the sentence, the two choice words were printed in capital letters and in a bigger font size (20 pt) than the other words of the sentences (14 pt).

We gave the children three practice items, where the choice was between two words that only differed from each other by one phoneme (and one letter). We did not want to use geminates in the practice words, since this might have given the children some guidelines about how to perform the task. However, the words had to differ from each other both orthographically and phonologically. Therefore, we used letters that were visually confusable, e.g. N-M, R-P and L-T.

2.1.2.2 Procedure

The experimenter started by introducing herself and a toy cat she had with her. Then, she handed out the response sheets and children heard the following instructions: The cat had found some new words nobody had ever heard before. The problem was that now someone should decide how to spell these words, but the cat, unlike all the children, did not know anything about reading or spelling of words. Therefore, he needed help and was asking the children to tell how the words should be spelled. The experimenter would help here and read out and repeat every row of words that the children had on the sheet in front of them. The children's task was "to circle one of the words written with big letters, the one that looked more like a real word".

Any questions the children had were answered and the experiment started with the three practice items. Between these, the experimenter went round the classroom making sure every child was only circling one word at a time, not more and not less, and in general knew what to do.

2.2 Results

2.2.1 Scoring

Children received one point every time they chose a word with a geminate in it. Thus, we had a measure of how likely they were to choose a geminate in the various experimental conditions, and by looking at the condition in question, we could determine whether the choice was appropriate orthographically, phonologically or both. The means of the scores in the different conditions are presented in Figure 1. It is important to note that the only time that the choice of a geminate was appropriate was in the Doublet/Medial condition.

INSERT FIGURE 1 ABOUT HERE

The scores showed that children in Year 1 chose initial geminates less often than medial geminates, irrespective of whether the phoneme that they heard was short or long. Year 2 children chose initial geminates less often than Year 1 children, and they were also unaffected by the length of the initial phoneme they heard. Year 3 children were even less likely to choose initial geminates than the younger children. However, Year 1 children chose medial geminates upon hearing a long phoneme only slightly more often than when hearing a short phoneme, and their scores were very close to the chance level of responding, which was 3. Year 2 children chose geminates more often for long medial phonemes, and therefore seemed to pay more attention to the length of medial phonemes that they heard. The third-year children were almost always choosing a medial geminate when they heard a long medial phoneme and a single letter when hearing a short medial phoneme.

These results were confirmed by repeated-measures Analyses of Variance. Therefore, we can say that although first-year children were more likely to accept geminates in the medial position

than the initial position, whether they heard a long or a short phoneme in either position did not make a significant difference to their answers. Second-year children were significantly more likely to choose geminates when they heard a long phoneme than when they heard a short one, but this was true only when the long phoneme was in the medial position. The second-year children rejected doublets in the initial position irrespective of whether they heard a long or a short phoneme. Hence, we did not find an interaction between orthographic and phonological knowledge, where children would have been more likely to accept initial geminates because they pay attention to the length of the phoneme involved. The third-year children were even more likely than the second-year children to choose medial doublets when they heard a long phoneme, and less likely to do so when they heard a short phoneme. Also, they almost always rejected initial geminates. Consequently, it appears that third-year children pay attention to the length of the phoneme and are able to represent it correctly in their spelling, as well as following the orthographic rule that word-initial geminates are illegal.

We used one-sample t-tests to see whether children's scores were significantly different from chance. These showed that the second- and third-year children's scores were significantly different from chance in all conditions. However, the first-year children's scores in the medial conditions were not significantly different from chance, although they were in the correct direction. This further supports our suggestion that the first-year children did not pay attention to length of the phonemes they heard. If they had taken length into account, they should have performed better than chance in the task.

2.3 Discussion

The results show us that first-year children already have some knowledge of the fact that words are not supposed to begin with a geminate. However, these young children do not pay attention to the length of phonemes yet, and therefore they are equally likely to choose a medial

geminate whether they hear a long or a short phoneme. Second-year children have enough experience of reading and spelling to know that phoneme length matters, and they also know that consonant length is represented with geminates. Therefore they are significantly more likely than chance to choose medial geminates when they hear long phonemes and single letter spellings when they hear a short phoneme in the middle of the word. However, although the second-year children pay attention to length most of the time, they are far from perfect, because they choose geminates for long phonemes about two thirds of the time and a geminate for short phonemes about one third of the time. In contrast, third-year children are almost at ceiling in choosing a geminate for long word-medial phonemes and a single letter spelling for short word-medial phonemes. Therefore, we can say that by the beginning of their third school year, children have mastered both the orthographic and phonological rules governing the use of geminates.

It is important to note that when taking the length of phonemes into account, second- and third-year children also follow the orthographic rule concerning initial geminates; they are significantly less likely than the first-year children to choose initial geminates. This indicates two issues. Firstly, children's orthographic knowledge develops from the first year to the second year, and older children are overall less likely to choose initial geminates. Second-year children are not significantly different from third-year children in this skill, so it seems that this orthographic rule is mastered during the first year of schooling, although, as we pointed out, school entrants already have some knowledge of it. Secondly, phonological information does not override orthographic information, and thus learning about the length of phonemes during the first school year does not increase children's tendency to choose initial geminates. In general, the condition where orthographic and phonological knowledge conflict does not produce any additional problems to children in any age group.

3 Study 2

Study 1 showed us that children have access to orthographic knowledge from early on, and also begin to learn about the phonological function of geminates during the first school year. This indicates that children are able to use multiple kinds of information to work out the rules governing the Finnish spelling system. Another possible source of information in spelling is morphology. Ravid & Gillis (in press) have suggested that a language that is morphologically complex predisposes children to use their morphological knowledge in spelling. They compared Hebrew (a morphologically complex language with a deep orthography) and Dutch (a morphologically simple language with a relatively transparent orthography) and found that children were using morphological information in spelling more extensively in Hebrew than in Dutch. Nunes, Bryant and Bindman (1997) have showed that English children's morphological awareness gradually begins to affect their spelling as they develop more sophisticated spelling strategies and start to appreciate the morphological regularities of the English spelling system. Similar results have been demonstrated in Portuguese and Greek (Rego, 1999; Bryant, Nunes & Aidinis, 1999). This work prompts us to ask whether morphological information might also be involved in Finnish children's spelling attempts in addition to the influences of orthographic and phonological information that we already saw. Since Finnish is morphologically complex, it should present numerous opportunities for the use of morphological information.

Investigation of children's geminate spelling provided us a way of inspecting this question. We asked children to do a spelling-to-dictation task, where we compared spelling of geminates such as *ss* and *ll* when these appear either in the word stem or in the case inflection. We also included other types of consonant clusters, *st* and *lt*, to get a measure of children's spelling of consonants when it is only necessary to consider phoneme identity, not phoneme length. We hypothesised that if children were able to take advantage of the morphological structure of the word, they would be likely to spell geminates and consonant clusters better when these are in case inflections than when

they are in word stems. Several studies reviewed earlier suggest that speakers of highly inflected languages parse inflected words into stems and inflections (Caramazza *et al.*, 1988; Niemi *et al.*, 1994). If this is the case, children would be likely to find spelling of consonant sequences easier in inflections. Case inflections have low type frequency as there are only 14 of them, while their token frequency is high, because they are repeatedly used both in spoken and written language. In contrast, word stems have high type frequency, since there are so many different stems, but their token frequency is low due to the low number of stems of any one type. Therefore, case inflections lend themselves much better to frequency-based learning than stems do. We decided to test the children twice, halfway through and at the end of their first school year, to be able to find out whether there is a developmental pattern to children's possible use of morphological information in spelling.

3.1 Method

3.1.1 Participants

We tested 41 children in a primary school in Espoo, Finland, in December of their first school year and four months later in April, towards the end of their first school year. The mean age of the children at Time 1 was 7;4 years and at Time 2 7;8 years. There were 15 boys and 26 girls.

3.1.2 The task and procedure

The spelling tasks discussed here were given as part of a larger longitudinal study of children's spelling development, and were administered in a single session to groups of about ten children each.

The Spelling task consisted of 24 nouns that varied according to two variables: position of phoneme cluster (Word Stem/Case Inflection) and type of phoneme cluster (Geminate/Mixed Cluster). The design of the tasks is presented in Table 5. We had to use a different set of words at

Times 1 and 2, since spelling development in Finnish is such a rapid process that we were concerned about possible ceiling effects at Time 2 if we had used the same word set. This time there was no conflict of different kinds of information, as in Study 1. Instead, we were interested in whether the morphological structure of the words would affect children's spelling performance. The reliability of the Spelling task at Time 1 was .92 and at Time 2 it was .87 (Cronbach's alpha).

INSERT TABLE 5 ABOUT HERE

3.2 Results

3.2.1 Scoring

We gave children one error point for each incorrect spelling of a cluster and scored each condition separately. As a result, each child got four error scores. The means and standard deviations of these are presented in Tables 6 (Time 1) and 7 (Time 2).

INSERT TABLES 6 AND 7 ABOUT HERE

We can see that it made no difference whether the geminates were in the word stems or inflections at Time 1. However, at Time 2, geminates produced much higher error rates in the Stem condition than in the Inflection condition, while there was no corresponding difference in the Mixed Cluster condition. In general, geminates seem to cause more errors than mixed clusters whether they were in the Stem or the Inflection condition. Children's most common error in geminate spelling was to leave out one letter of the geminate. Mixed clusters caused children to spell letters incorrectly and to miss out letters. These types of errors occurred both with the first and the second letter of the cluster.

3.2.2 Analyses

We inspected the means by a repeated-measures Analysis of Variance, where the within-subjects variables were Position (Stem/Inflection) and Phoneme (Geminate/Mixed Cluster).

3.2.2.1 Time 1

At Time 1, we found a significant main effect of Phoneme ($F(1,40)= 25.128, p< .001$), because children made more errors when they spelled geminates than when they spelled mixed clusters. Our previous results led us to expect this, and the difference is particularly dramatic at this early phase of spelling development, when many children largely ignore phoneme length instead of just making the occasional error with geminates. There was no significant main effect of Position or any interactions, so it seems that children's spelling at the end of their first term at school is not facilitated by the morphological structure of the word at all.

3.2.2.2 Time 2

At Time 2, we found significant main effects of both Position ($F(1,40)= 13.892, p= .001$) and Phoneme ($F(1,40)= 20.811, p< .001$) and a significant interaction between these ($F(1,40)= 22.732, p< .001$). We inspected the interaction by paired-sample t-tests with Bonferroni corrections ($\alpha= .013$). These indicated that while geminates caused significantly more errors in stems than in case inflections ($t(40)= 5.294, p< .001$), it made no significant difference whether mixed clusters appeared in stems or inflections. This suggests that the morphological structure of words influences children's spelling and supports our hypothesis about the nature of this facilitation; children are spelling geminates better in case inflections than word stems. What is interesting is that facilitation only occurs in geminate spelling, which agrees with the idea that geminate spelling requires somewhat different processes than mixed cluster spelling. Mixed clusters, which only require sound-letter decoding, are quite straightforward to deal with in a transparent orthography such as Finnish. Children's error rates in spelling these types of clusters are relatively low, and therefore there is not much room for improvement through other strategies. In contrast, geminate spelling requires children to pay attention to both phoneme identity and phoneme length. This is more demanding than just sound-letter decoding and therefore additional strategies, such as cues provided

by the morphological structure of the word, are more likely to aid geminate than mixed cluster spelling.

4 General Discussion

These two experiments set out to examine how children's orthographic, phonological and morphological knowledge affect their developing spelling skills. The particular aspect of spelling that we were interested in was geminate spelling, which is interesting for two reasons. Firstly, the phonological aspect of the geminate spelling rule in Finnish determines a relationship between phoneme length and geminates, since geminates are used to spell long phonemes. Secondly, the geminate spelling rule in Finnish is conditional on orthographic position, since word-initial geminates are not allowed even when a long consonant appears at the beginning of the word.

Study 1 showed that already first-year children at the beginning of the school year are less likely to accept geminates in the initial than in the medial position of the word. Therefore, they demonstrate some knowledge of the orthographic rules of Finnish. However, first-year children are largely ignorant of the phonological function of geminates. Consequently, they are equally likely to accept word-medial geminates and single letter spellings whether they hear a short or a long consonant in the middle of the word.

Second-year children, on the other hand, reject word-initial geminates even more readily than first-year children and are also able to pay attention to the length of the phoneme they hear. Because of this, they choose medial geminates predominantly when they hear a long phoneme and single letter spellings upon hearing a short phoneme. This tendency is even more pronounced in the third-year children's performance, which is close to ceiling in all conditions. The conflict of phonological and orthographic information in the word-initial geminate condition does not affect the second- or third-year children. Although they do know that long phonemes should be spelled with geminates, they do not accept initial geminates even when they hear a long phoneme in the initial position of the word.

Study 2 investigated whether children benefit from the morphological structure of the word when they spell geminates. The results showed that the morphological structure of the word makes no difference in children's performance halfway through the first school year, but already four months later in the spring these same children were spelling geminates significantly better in case inflections than word stems. Therefore, the morphological structure of the word does facilitate spelling, although only of geminates. Mixed consonant clusters were spelled equally well whether they were in stems or inflections.

The results of Study 1 show that even children who have received no formal instruction in literacy skills can have some knowledge of what words are supposed to look like. Before entering school, children will have seen written words on many occasions and are able to notice some characteristics of words. Given the late age of school entry in Finland, children's opportunities for this type of informal learning are many. Although the first-year children still chose an initial geminate about one third of the time, the fact that they were better than expected by chance is nevertheless quite a sophisticated achievement, since the Finnish orthography does allow word-initial vowel doublets. Therefore, children cannot merely reject all word-initial letter doublets. Rather, they have to have some idea of the doublets that are allowed and those that are not. As Pacton *et al.* (2001) pointed out, French first-year children are more likely to reject double vowels, which are not legal in French, than double consonants, which are. This agrees with the results of Finnish children. One possible way for acquiring this kind of knowledge is through frequency-based learning, which has been suggested both by Pacton *et al.* (2001) and Cassar & Treiman (1997). However, we would have to conduct more carefully controlled experiments to test this hypothesis in Finnish children.

It is not surprising that the first-year children at the beginning of the school year in Study 1 were unaffected by phoneme length. Although this is a distinction that children of this age do not find particularly hard in spoken language, the concept of short and long phonemes only seems to

become explicit in reading and spelling gradually during the first school year. This was confirmed by the results of the older children. They were able to choose geminates appropriately according to whether they heard a long or a short phoneme and this skill improved with increasing age.

An interesting finding was that although second- and third-year children had conflicting orthographic and phonological information in the condition where they heard a long initial phoneme, this did not affect their performance. Thus, children take into account what they know about the orthographic constraints of their language when they begin to work out the phonological function of geminates. The independence of orthographic and phonological information is illustrated by the fact that during the year when children learn to represent long phonemes with geminates, they reach near floor level in the number of chosen geminates in the initial position of the word. However, it is possible that when children start learning the phonological function of geminates, there is a time when they also accept initial geminates, phonological information overriding the orthographic rules about geminates. We might not have caught this phase in our study, since detecting it would require testing children longitudinally and probably also at shorter intervals than a whole year. Additional studies are thus required.

Therefore, during the first year of schooling children's orthographic knowledge of geminates solidifies and they also start to appreciate the phonological function of geminates. This is also the time when, as shown by Study 2, children become sensitive to the morphological structure of words, specifically in geminate spelling. Halfway through the first school year, many children are still likely to decode words laboriously, using sound-to-letter mappings and failing to pay attention to geminates or the morphological structure of the words they are spelling. However, just a few months later we see significant facilitation of geminate spelling in case inflections. This suggests that geminate spelling requires different processing than other consonant clusters, which can be spelled more easily by using the regular, one-to-one mappings between phonemes and graphemes that Finnish provides. The spelling of geminates, on the other hand, is a more complicated process

and is affected by children's ability to analyse the morphological structure of the word they are spelling.

Why do children spell geminates better in case inflections than in word stems? There are two potential explanations. One possibility is that these children, like the English children studied by Nunes *et al.* (1997), begin to notice the morphological function of inflections because of their developing morphological awareness skills. This would require further study looking at whether children's performance in morphological awareness tasks would predict their ability to spell case inflections better than their spelling of stems.

It is also possible that the results are due to more implicit morphological processing, i.e. the ability of speakers of highly inflected languages to parse words to stems and inflections. Additionally, as we already suggested, the high token frequency and low type frequency of case inflections make them salient learning targets, while stems possess the opposite frequency characteristics. Children's performance in the orthographic constraints test suggested that they are sensitive to the position of geminates in words from early on; this could in part explain their superior spelling of geminates in case inflections. Our present results do not allow us to determine in more detail whether morphological parsing is involved, which would necessitate a degree of morphological knowledge, or whether children just pay attention to case inflections because they are frequent orthographic patterns.

However, independent of whether the difference between spelling geminates in stems and inflections is due to morphological awareness or learning of orthographic spelling patterns, the phenomenon demonstrates that as early as in the second half of their first school year, children begin to appreciate the phonological aspect of the geminate spelling rule. Moreover, children pay attention to spelling patterns in different parts of words and may possibly parse words to stems and inflections as they spell them. The finding that the difference between spelling stems and inflections only emerges for geminates, not mixed consonant clusters suggests that it is possible for children to

learn the consistent phoneme-grapheme mappings without the help provided by the morphological knowledge. Children learn these mappings before they start to appreciate the morphological structure of words. However, learning to spell geminates takes longer, and improves with children's growing ability to pay attention to the morphological structure of words.

In summary, the studies we have presented suggest that Finnish children can use multiple sources of information when they are learning to spell. Even before the start of formal teaching, first year children know something about the legal position of geminates, and during their first year at school they begin to appreciate the phonological function of geminates, i.e. that they represent phoneme length. During this process, children seem to be helped by the morphological structure of words, since they are more likely to spell geminates correctly in case inflections than word stems. Consequently, our results do not support strict stage models of spelling, which postulate that children in different stages of spelling development use qualitatively different information in spelling and that e.g. orthographic information in spelling would be a sophisticated strategy which children in the semiphonetic or phonetic stage of spelling would not be able to consider. In contrast, children seem to be able to use different kinds of knowledge to their advantage in spelling from early on.

One issue that we need to address is the way in which these results might be important for the study of spelling in general. As we already pointed out earlier, English is quite an exceptional orthography, thanks to its deep morphophonological structure. Despite this, most of spelling models are based on the English orthography. Studies of orthographies other than English are trying to extend the work done in English by answering two questions. Firstly, they attempt to find out which processes affect spelling development independently of the structure of orthography. Secondly, they try to investigate how properties characteristic to particular orthographies affect the ways in which children learn to spell.

The two studies presented here contribute to the first question by showing that just as English and French children, young Finnish children seem to be sensitive to the legal position of letter doublets. Also, like English children, Finnish children take longer to learn the phonological than the orthographic aspect of the geminate spelling rule. With regards to the second question, the most obvious characteristics specific to Finnish is that it requires children to learn the difference between short and long phonemes in spelling. This suggests that learning about the phonological properties of words is not just about learning phoneme identity and how this is expressed in spelling, but other phonological aspects can be involved as well, depending on the language in question. Several other languages code for phoneme length, for example in Hungarian this is done by using accents on top of letters. Also, it is important to note that spelling development in Finnish is much more rapid than in English, and therefore children start appreciating the phonological role of geminates already during the second half of their first year at school, while English-speaking children only learn it in the sixth grade.

There are several ways in which we could improve the studies to investigate the conclusions even further. Firstly, it would be important to test the children's reading and spelling levels to see how this relates to their ability to use orthographic and phonological knowledge in the orthographic constraints test. It would also be informative to do a version of the orthographic constraints test with vowel doublets as well as geminates as targets, since Finnish does allow vowel doublets in word-initial positions. This way we could find out whether children are merely using a general rule of the type "no doublets at the beginning of words allowed" or whether they are actually sensitive to the different properties of consonants and vowels. Thirdly, we need to explore the reason for the facilitation we find in inflections as opposed to word stems. To investigate whether it is due to morphological processing, we should look at children's performance in morphological awareness tests. A way to find out if the facilitation is due to orthographic processing of the position of case inflections would be to ask children to spell words where the case inflection is not the final

inflection of the word. This way, we would gain more detailed information about the kinds of knowledge young children can use to their advantage when they are learning to spell.

Acknowledgements

I would like to thank Peter Bryant for his comments on an earlier version of this chapter and many helpful discussions. I am also grateful to the teachers and pupils of the schools in Finland who participated in the studies presented here.

References

- Bryant, P.E., Nunes, T. & Aidinis, A. (1999). Different morphemes, same spelling problems: Cross-linguistic developmental studies. In M. Harris & G. Hatano (Eds.), *Learning to Read and Write: A Cross-linguistic Perspective* (pp. 112-133). Cambridge: Cambridge University Press.
- Cassar, M. & Treiman, R. (1997). The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *Journal of Educational Psychology*, 89, 631-644.
- Caramazza, A., Laudana, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, 28, 297-332.
- Carney, E. (1994). *A Survey of English Spelling*. London: Routledge.
- Ehri, L. (1992). Review and commentary: Stages of spelling development. In S. Templeton & D. Bear (Eds.), *Development of orthographic knowledge and the foundations of literacy: A memorial Festschrift for Edmund H. Henderson*, (pp. 307-332). Hillsdale, NJ: Lawrence Erlbaum.
- Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson, M. Coltheart and J. Marshall (Eds.), *Surface Dyslexia*. London: Lawrence Erlbaum Associates.
- Laaksonen, K. & Lieko, A. (1998). *Suomen kielen äänne- ja muoto-oppi* [The phonology and structure of the Finnish language]. Helsinki: Finn Lectura.
- Lehtonen, J. (1970). Aspects of Quantity in Standard Finnish. *Studia Philologica Jyväskyläensia*: 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. Berninger (Ed.), *The varieties of orthographic knowledge II: Relationships*

to phonology, reading, and writing, (pp. 177-204). Netherlands: Kluwer Academic Publishers.

Niemi, J. & Laine, M. (1995). Suomen kielen sanojen mentaalisen käsittelyn malli [A mental model for the processing of Finnish words]. *Suomen Logopedis-foniatrinen aikakauslehti*, 14, 65-71.

Niemi, J., Laine, M., & Tuominen, J. (1994). Cognitive morphology in Finnish. Foundations of a new model. *Language and Cognitive Processes*, 9, 423-446.

Nunes, T., Bryant, P. & Bindman, M. (1997). Morphological spelling strategies: Developmental stages and processes. *Developmental Psychology*, 4: 637-649.

Pacton, S., Perruchet, P., Fayol, M., & Cleeremans, A. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General*, 130, 401-426.

Ravid, D. & Gillis, S. (in press). Teachers' perception of spelling patterns and children's spelling errors: A cross-linguistic perspective. In M. Neef, A. Neijt & R. Sproat (Eds.), *Consistency in writing systems*. Tübingen: Niemeyer Verlag.

Rego, L. (1999). Phonological awareness, syntactic awareness and learning to read and spell in Brazilian Portuguese. In M. Harris & G. Hatano (Eds.), *Learning to Read and Write: A Cross-linguistic Perspective* (pp. 71-88). Cambridge: Cambridge University Press.

Treiman, R. (1993). *Beginning to Spell*. New York: Oxford University Press.

Table 1 Examples of how length affects meaning in Finnish.

Finnish	English
<i>tak<u>a</u></i>	back
<i>tak<u>ka</u></i>	a fireplace
<i>taak<u>ka</u></i>	a burden
<i>tak<u>aa</u></i>	from behind
<i>tak<u>kaa</u></i>	of a fireplace
<i>taak<u>kaa</u></i>	of a burden

Table 2 Examples of Finnish morphology.

Finnish	Translation
TALOISSAMMEKIN	in our houses as well
TALO + <i>I</i> + SSA + <u>MME</u> + <u>KIN</u>	house + <i>s</i> + in + <u>our</u> + <u>as well</u>

Table 3 Some Finnish case inflections showing how the stem of the word changes (stem of the word in bold).

Finnish	English
vesi	water
veden	of water
vettä	some water
vetenä	as water
vedeksi	into water
vedessä	in water
vedestä	from water
veteen	to water
vedellä	with water

Table 4 The design of the orthographic constraints task and examples of items (wrong choices in parentheses).

		Position	
		Initial	Medial
Phoneme	Short	Paksu Soaki (SSoaki) painaa.	Kaunis neiLe (neiLLe) laulaa.
		Nuori Laire (LLaire) nukkuu.	Vanha raiSi (raiSSi) lepää.
	Long	Heitä Kitri (KKitri) ulos.	Pieni muoSSi (muoSi) nauraa.
		Kisko Patso (PPatso) irti.	Vihreä taiLLe (taiLe) herää.

Table 5 The design of and some examples of the words in the Spelling Task.

		Position	
		Word Stem	Case Inflection
Phoneme	Geminate	reiSSu	kerhoSSa
		siveLLin	tunniLLa
	Mixed Cluster	riSTeys	kuoroSTa
		keLTainen	hiekaLTa

Table 6 The means (and standard deviations) of the children's error scores in the Spelling Task at Time 1 (maximum score 6 in each condition).

	Geminate	Mixed Cluster
Word Stem	1.80 (2.08)	.20 (.51)
Case Inflection	1.88 (2.52)	.46 (.78)

Table 7 The means (and standard deviations) of the children's error scores in the Spelling Task at Time 2 (maximum score 6 in each condition).

	Geminate	Mixed Cluster
Word Stem	2.56 (2.21)	.56 (.98)
Case Inflection	1.34 (1.98)	.61 (.92)

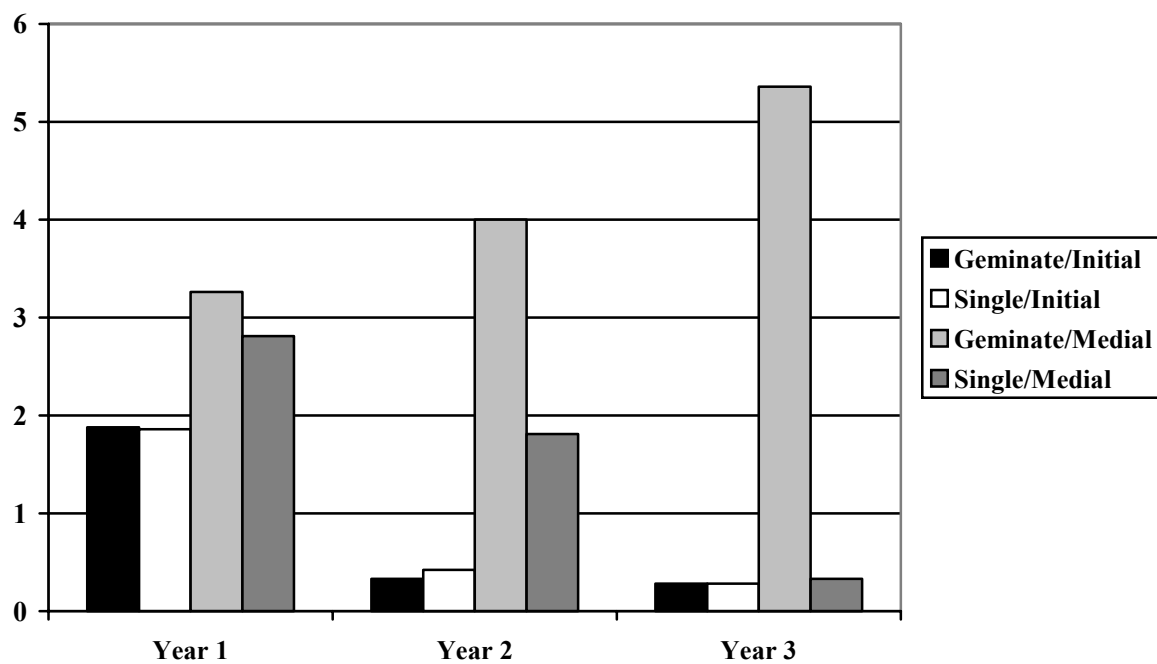


Figure 1 The mean number of geminate choices in different conditions (out of 6).