Available online at www.sciencedirect.com
SCIENCE DIRECT®
Brain and Language 90 (2004) 338–352
Brain and Language
www.elsevier.com/locate/b&l

Segmenting two-phoneme syllables: Developmental differences in relation with early reading skills☆

Astrid Geudens,[a,*] Dominiek Sandra,[a] and Wim Van den Broeck[b]

[a] *University of Antwerp—UFSIA, Center for Psycholinguistics, Antwerp 2000, Belgium*
[b] *University of Leiden, The Netherlands*

Accepted 3 December 2003
Available online 4 February 2004

Abstract

This study explored developmental differences in children's segmentation skills of VC and CV syllables (e.g., /af/ and /fa/) in relation to their early reading abilities. To this end, we followed a subgroup of Dutch speaking prereaders who participated in Geudens and Sandra (2003, Experiment 1), and replicated the segmentation task in first grade, at the outset of phonics reading instruction. Reading abilities were assessed after 6 and 9 months. First, we confirmed that VCs offer an easier context to isolate phonemes than CVs. Second, matching analyses showed that this development from VC to CV segmentation posed comparatively increasing difficulties for poor segmenters. Third, this qualitatively different development was reflected in early reading performance. Our data emphasize the importance of phonetic factors and instruction-based experiences in phonological development.

Keywords: Segmentation development; Explicit phoneme awareness; Phonetics; Individual differences; Reading; Necessity; Sufficiency

1. Introduction

One of the basic steps in learning to read and spell in an alphabetic script is to understand, consciously, that spoken words can be broken up into separate sounds, not only at the level of larger units like syllables, and consonant–vowel combinations, but especially at the most fine-grained phoneme level (Adams, 1990; Caravolas, Hulme, & Snowling, 2001; Ehri, 1998; Morais, Mousty, & Kolinsky, 1998; Nation & Hulme, 1997; Share, 1995). This *explicit* awareness of phonemes does not emerge spontaneously: phonemes are abstract linguistic representations of sounds and do not exist as discrete elements in the rapidly paced human speech signal (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). Hence, in order to realize that /pen/ consists of the phonemes /p/, /e/, /n/, the child must find a code to break through the continuous airflow as well as push down from the level of surface phonetic segments, and such code is not provided outside the context of literacy (Morais, Bertelson, Cary, & Alegria, 1986). Only by learning the letters that symbolize the short-lived sounds, the child is able to abstract from phonetic variation and build up precise and consciously accessible representations of the syllables' phonemes (see also Adams, 1990; Hohn & Ehri, 1983; Johnston, Anderson, & Holligan, 1996). At the same time, phonemic awareness—developing in interaction with print experience—helps to facilitate the process of mastering simple and more complex grapheme–phoneme conversion rules (Morais et al., 1998; Perfetti, Beck, Bell, & Hughes, 1987). Although it has become increasingly clear that training of "pure" phonological skills without instruction of letters generally produces very modest gains in reading and spelling (see Krashen, 2001; Van den Broeck, 1997), an interactive training that combines both phonology and orthography seems highly successful. Many studies have demonstrated that training children on phoneme segmentation skills in combination with teaching of grapheme–phoneme

☆ The name order of the second and third author is alphabetical (equal contribution). We are most grateful to all of the children, teachers, and headteachers of the primary schools *Reuzepas* en *Delta* (Oud-Turnhout) who cooperated in and supported this study so generously.
* Corresponding author. Fax: +32-3-220-4269.
E-mail address: astrid.geudens@ua.ac.be (A. Geudens).

0093-934X/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/S0093-934X(03)00446-2

correspondences (GPCs) is one of the best ways to promote their reading and spelling skills (e.g., Ball & Blachman, 1991; Byrne & Fielding-Barnsley, 1991; Hatcher, Hulme, & Ellis, 1994; Hohn & Ehri, 1983).

In this study, we focused on individual differences in Dutch-speaking (Flemish[1]) children's early development of phoneme segmentation skills. Even though most children who are learning to read in this relatively consistent orthography (Martensen, Maris, & Dijkstra, 2000) generally achieve high scores on segmentation tasks rather quickly (Patel, 2001; Reitsma, 2002; cf. Landerl & Wimmer, 2000; in German), some children encounter problems in mastering this skill (cf. Patel, 2001; Van der Wissel, 1984). For these children, it is important to determine what the factors are that hamper the ability to break up syllables intentionally so that reading instruction can develop from the easiest contexts to more difficult ones. In a previous paper (Geudens & Sandra, 2003), we studied the factors that influence prereaders' and beginning readers' early segmentation development of simple CVs, VCs, and CVCs. Several researchers have suggested that children find it harder to segment VC than CV strings because breaking up a VC entails the division of a linguistic subsyllabic unit, the *rime*[2] (e.g., Gombert, 1996; Schreuder & Van Bon, 1989; Stahl & Murray, 1998; Treiman, 1992; Van Bon & Duighuisen, 1995). However, in contrast, we could not find any effect that indicates the importance of the phonological onset-rime structure in explicit awareness. Rather than being determined by a fixed linguistic structure, the children's segmentation patterns seemed to fluctuate depending on articulatory- and perceptual-phonetic factors.

These phonetic factors also enabled us to account for opposite cohesion patterns reported in the literature on phonological awareness (e.g., Bus, 1985; Coleman, 1970; Lewkowicz & Low, 1979; Martinot & Gombert, 1996; Uhry & Ehri, 1999). For instance, we observed that children's difficulties to break up the speech signal increased as a function of the vowel-likeness (sonority) of consonants preceding and following the vowel. Therefore, the treatment of the sonority factor as a random variable in earlier segmentation experiments may have caused problems, especially when this resulted in the inclusion of more sonorants in postvocalic than prevocalic position (see also Duighuisen, Van Bon, & Schreuder, 1991). This could be especially true for languages like English and northern standard Dutch, where sonorants like /l/ and /r/ are vocalized in postvocalic position, making these phonemes even harder to discriminate from the resembling vowel sound (Collins & Mees, 1981; Sproat & Fujimura, 1993; Van Reenen & Jongkind, 2000). When controlling for the sonority factor as much as possible in our study with children speaking the southern variant of standard Dutch, we observed a significant advantage of VC over CV strings in two-phoneme and three-phoneme syllables. Letter knowledge could not have confounded the effect since initial and final consonants were the same (e.g., /at/ vs. /ta/). We concluded that a VC offers the best phonetic conditions for separating sounds in the acoustic signal, at least in southern Dutch, and that children's early segmentation abilities develop from VCs with non-sonorous consonants to CVs.

Following the rationale that explicit awareness requires breaking through the continuous speech flow, it is not surprising that early segmentation skill is strongly influenced by perception and articulation.[3] This leads to the prediction that poor segmenters may have particular difficulties to abstract from surface phonetic features. Besides opportunities to learn about sounds and letters at home, a lower degree of phonological sensitivity may restrain them from acquiring GPCs (Bowey, 1994). Correspondingly, it may be harder for them to realize that different pronunciations of a particular sound (i.e., allophones) are in fact realizations of the same phoneme. This fuzziness may in its turn hamper the development of refined phonemic and graphemic representations.

These observations and reflections inspired us to explore three central issues. In the first part of our study, we followed a subgroup of prereaders who participated in our previous segmentation task in June (Geudens & Sandra, 2003, Experiment 1) and examined whether the VC advantage could still be replicated at the outset of phonics reading instruction, in September. At this time, the children had received explicit instruction about the correspondence between graphemes and phonemes in blending and segmentation tasks. In the second part of our study, we relied on the follow-up data to test our prediction that the more complex phonetic context in CV strings poses comparatively more segmentation problems for poor segmenters. If phonological awareness is influenced by phonetic factors and if poor

[1] Flemish is a southern variant of standard Dutch as spoken in northern Belgium, and mainly differs from the northern variant in the Netherlands in terms of pronunciation characteristics.

[2] Several linguists have proposed that the basic syllable structure is best represented as a hierarchical onset-rime structure (e.g., Fudge, 1987; Trommelen, 1984; but see e.g., Pierrehumbert & Nair, 1995; Yip, 2002). In the psycholinguistic literature, the onset is conceived as the consonant or consonant group preceding the vowel (e.g., *sl* in *sleep*); the rime is composed of the vowel and the following consonant(s) (e.g., *eep* in *sleep*). The onset-rime distinction is considered to be a psychologically relevant structure, having important implications for phonological/reading development (e.g., Adams, 1990; Treiman, 1992).

[3] In this respect, the term "phonetic segment segmentation" may be a more proper alternative to refer to the task at this early stage than the term "phoneme segmentation" because the children's knowledge may concern phones rather than abstract phonemes (see also Content, Kolinsky, Morais, & Bertelson, 1986; Scholes, 1998).

segmenters experience particular difficulty to map phonetic properties onto phonemes, these poor segmenters should find it harder to progress from less difficult VC strings to more difficult CV strings compared to good segmenters. In order to test this hypothesis, we matched the VC scores of the poor segmenters in September to the VC scores of the good segmenters in June and then compared their CV scores. In the third and final part of this paper, we studied how the children's poorer performance on the segmentation task was related to their early reading abilities. Therefore we administered standardized reading tests in February and June in the children's first year of reading instruction.

2. Follow-up of the segmentation study in first grade

In our previous study (Geudens & Sandra, 2003; Experiment 1), we explored prereaders' ($N = 56$) performance on a segmentation task of reversed CVs and VCs. At the time of testing in June, no exercises in segmentation or reading were part of the curriculum, but almost all children knew a few letters, mostly from their first name. The present follow-up experiment was performed three months later in September with a subsample of these children ($n = 41$). ANOVAs with Subsample as an extra between-participants factor confirmed that this sample was representative of the full sample from the earlier study (F's < 1). Besides exploring children's segmentation skills, we also administered a letter sound knowledge test and followed the children's development of reading proficiency.

2.1. Method

2.1.1. Participants

A group of 41 first-graders took part in this study (mean age = 6.3 years, range = 6.2–7). They were attending two regular primary schools in Flanders (northern Belgium) and had already participated in the same task three months earlier. The children were native speakers of southern Dutch and had no documented hearing or speech difficulties. At the time of testing in September, the curriculum included explicit phonics training (*Leeslijn* [Reading-Line], De Baar, 1995). For each new word, letters were presented and corresponding letter sounds (pronounced as schwa names) were identified. GPCs were practiced in other words as well, each time segmenting the phonemes while keeping track of the graphemes, and blending them afterwards.

2.1.2. Materials and design

2.1.2.1. Segmentation task. Materials (Appendix A) consisted of 32 pairs of CVs and reversed VCs (Phoneme Order). All items were phonologically legal. Items were composed from a set of seven phonologically long[4] vowels (/o/, /a/, /e/, /i/, /y/, /u/, /ø/) and one diphthong (ʔœyʔ) and a set of consonants representing four sonority levels (Consonant Type): stop /p/ or /t/, fricative /f/ or /s/, nasal /m/ or /n/, and liquid /l/ or /r/. Each vowel was combined with four consonants, one of each consonant type. Given our purpose to use items that were each other's reversal (e.g., /to/ and /ot/), we only used consonants that could occur both in initial and final position. Thus, each vowel produced four CVs and four VCs, yielding 32 items of each type (8 vowels × 4 consonant types). Our aim was to construct pseudowords only, but the restrictions on item selection forced us to include words as well (43 pseudowords, 21 words).[5] Two lists (List) were composed, each comprising 16 CVs and 16 VCs. For one list, 16 CVs were derived from four vowels (4 vowels × 1 consonant of each type) and 16 VCs were derived from the remaining four vowels and consonants. The other list was obtained by reversing the items (in the same position). In addition, eight CVC pseudowords (e.g., /kɪg/) and their reversals (e.g., /gɪk/) were used as distractors. Each list was presented to one half of the participants and presented in two fixed pseudo-random orders, avoiding as much as possible successive trials sharing any phonemes. As the correlation of a test with itself is necessarily higher than the correlation with any other test, the intercorrelation between the VC and CV segmentation scores was taken as a conservative estimate of the test-reliability. The intercorrelation was .83, indicating good reliability.

2.1.2.2. Letter sound knowledge. The untimed letter sound knowledge test contained 26 single letters and 12 digraphs ($N = 38$), printed in the font of the children's reading books. As demonstrated by Hulsen and Voets (2003), test–retest reliability of this measure is good, .86.

2.1.2.3. Reading. In order to assess the children's reading skill we administered the standardized Cito reading test (*Cito Drie-minuten-test*, [Cito Three-minute-test]; Verhoeven, 1993) The Cito test is a timed reading test composed of three different lists of unrelated words which have to be read aloud as fast and as accurately as possible. For each separate list, the number of words read correctly in 1 min is the reading score. The first list consists of 150 monosyllabic words without consonant clusters. The second list consists of 150 monosyllabic words with consonant clusters. The third list is made up of 120 polysyllabic words, both with and without consonant clusters. In

[4] Phonologically long vowels differ in terms of phonetic duration: /o/, /a/, /e/, /ø/ are always phonetically long, whereas the high vowels /i/, /y/, /u/ are short in standard Dutch, except when directly preceding /r/ (Kager, 1989).

[5] An ANOVA treating Lexicality (word; pseudoword) as between-items factor showed that segmentation scores for words and pseudowords did not differ significantly ($F < 1$). The Lexicality by Item Type (CV; VC; CVC) interaction was not significant either $F = 1.49$, $p > .2$.

all lists a maximum variety of vowels and consonants is supported. Mean test reliabilities of the separate lists are .88, .94, and .91, respectively (Verhoeven, 1993).

2.1.3. Procedure

The segmentation task was administered after three weeks of phonics reading instruction in September. Children were familiar with the task because it resembled exercises in their curriculum and because of their earlier participation in the same test (Geudens & Sandra, 2003, Experiment 1). Instructions were along the following lines: "Which sounds do you hear in /ki/?". All practice trials (5 CVs and 5 VCs) and experimental items had been recorded in advance (by the first author) and were presented with a minidisc (Sony MZ-R37) and loudspeakers (Labtec Spin 60). For each trial the child had to listen very carefully, to repeat the syllable, and then articulate the sounds. Item repetition was used in order to verify the child's correct perception of the items. If the child did not repeat the item correctly, we presented it a second time (from minidisc). To make the scoring easier for the experimenter, the child had to clap hands while producing each sound. All errors were noted and all sessions were recorded on a second minidisc. Trials on which the child did not hear the stimulus correctly but performed a correct segmentation on the misperceived item (e.g., [oːl] is perceived as [oːw] and segmented as [oː]-[w]), were included in the analysis as correct segmentations (see Geudens & Sandra, 2003).

Immediately following the segmentation task in September, the letter sound knowledge test was administered. With the help of a cover sheet, each letter was presented individually. The score was made up of the amount of letter sounds pronounced correctly ($M = 11.17$; $SD = 7.25$).

The Cito reading tests were administered after six and nine months of reading instruction. In February (*Time 1*), after the children had been learning to read for six months, we administered the first list ($M = 24.68$; $SD = 14.82$). In June (*Time 2*), after nine months of reading instruction we again administered the first list ($M = 42.78$; $SD = 16.46$), and additionally measured the children's scores on the second and third word lists ($M = 25.22$; $SD = 15.05$ and $M = 15.32$; $SD = 11.12$, respectively). Children's reading scores at Time 2 were defined as the summed scores on the three reading lists in June ($M = 83.32$; $SD = 41.32$).

2.2. Results and discussion

2.2.1. Replication of the VC advantage

Our first aim in this study was to investigate whether the observed difference between VC and CV segmentation in our previous study in June could still be replicated in September, at the outset of phonics reading instruction. Table 1 shows the mean number of correct responses, percentages, and standard errors for each Phoneme Order and Consonant Type in the segmentation task. Repeated ANOVAs were performed for a $2 \times 4 \times 2$ (Phoneme Order $\times$ Consonant Type $\times$ List) design. Phoneme Order (CV; VC) and Consonant Type (stop; fricative; liquid; nasal) were within-participants and between-items factors. List (List 1; List 2) was a between-participants and between-items factor.

The main effect of Phoneme Order was significant, $F1(1,117) = 18.36$, $p < .0001$; $F2(1,48) = 68.66$, $p < .0001$, replicating the segmentation advantage of VCs over CVs in June. (The same results were obtained when we separately analyzed the data of those children who obtained zero-scores on the segmentation task in June ($n = 14$), $F1(1,36) = 13.44$, $p < .005$; $F2(1,48) = 36.21$, $p < .0001$.) We also explored whether the children's segmentation patterns would still be a function of the sonority type of the consonants. Sonority is generally defined as a ranking on a scale—from vowels, glides, liquids, nasals, to fricatives, affricates, stops—that is motivated by phonetic notions of perceptual salience (Goldsmith, 1990, p. 110–111). Sonorous consonants resemble vowels to a greater extent than obstruents because they involve a larger degree of openness of the vocal apparatus and are produced with a higher amount of energy relative to the articulatory effort (Goldsmith, 1990; Laver, 1994). Although sonority is generally considered a phonological notion (see Blevins, 1996), we were especially interested in the pronunciation of the different consonant types and hence approached sonority from a phonetic point of view. In June, the main effect of Consonant Type in the prereaders' data indicated that fricatives and stops (obstruents) and liquids and nasals (sonorants) behaved as two groups, vowel-like sonorants being harder to

Table 1
Mean number of correct responses (*M*), percentages, and standard errors (*SE*) in the two phoneme orders for all consonant types in the segmentation task

Phoneme order	Consonant type							
	Stop		Fricative		Nasal		Liquid	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
CV	2.45 (61%)	.25	2.85 (71%)	.25	2.47 (62%)	.28	2.70 (67%)	.25
VC	3.30 (82%)	.23	3.50 (88%)	.17	2.82 (70%)	.25	2.92 (73%)	.22

Note. The maximum mean score is 4.

separate from the vowel than obstruents. This pattern was especially characteristic for VCs and was not as clear-cut for CVs. A separate analysis of the CV data in June showed that prereaders actually found it as hard to segment CVs with stops as CVs with sonorants ($ps > .2$) and that CVs with fricatives were not only easier in comparison with sonorants ($ps < .05$), but also tended to be easier than items with stops ($ps < .06$). However, since the interaction effect between Phoneme Order and Consonant Type was not significant by items and only reached marginal significance by participants, we decided not to attribute theoretical importance to this finding in the prereaders' segmentation data. Yet, three months later in September, this pattern was consolidated. As in June, the ANOVA yielded a significant main effect of Consonant Type, $F1(3,117) = 8.40$, $p < .0001$; $F2(3,48) = 12.29$, $p < .0001$, but now, pairwise comparisons did not support an overall classification in terms of obstruents and sonorants. The effect of Consonant Type manifestly changed from VCs to CVs, as revealed by the significant Phoneme Order by Consonant Type interaction, $F1(3,117) = 7.28$, $p < .0001$; $F2(3,48) = 5.14$, $p < .005$. Posthoc analyses on the VC and CV data in September partly replicated the results from the separate analyses on the June data. In the case of VCs, pairwise comparisons were non-significant between stops and fricatives ($ps > .1$), and between nasals and liquids ($ps > .4$). All pairwise comparisons involving an obstruent vs. a sonorant were significant ($ps < .005$). Thus, in first grade as well as in kindergarten, children treated VCs with obstruents and sonorants differently. In the case of CVs, the pattern was different and not entirely compatible with our earlier findings (Geudens & Sandra, 2003). As in the previous experiment, stops created more difficulties than fricatives ($ps < .05$). Nasals were also more difficult than fricatives ($ps < .05$), but the other set of sonorants, i.e., liquids, were not ($ps > .1$). The difficulty of nasals did not differ from that posed by stops ($ps > .8$). Liquids were, surprisingly, easier than stops ($ps = .05$). Although we cannot offer an explanation for the behavior of sonorants in CVs, the entire data set clearly contradicts the received wisdom that the phonological onset-rime structure of the syllable governs children's abilities to segment syllables intentionally (see Geudens & Sandra, 2003).

In contrast, there were several indications in our experiment that phonetic factors rather than a fixed phonological structure determined children's segmentation patterns. For instance, the consonant's degree of vowel-likeness played an important role in VC segmentation as children found it harder to repeat and segment VCs with vowel-like sonorants than with obstruents (see also Coleman, 1970; Duighuisen et al., 1991; Geudens & Sandra, 2003; Yavas & Gogate, 1999). The results also seem to reflect articulatory-phonetic factors (see also Schreuder & Van Bon, 1989; Skjelfjord, 1987). As in Skjelfjord's (1987) segmentation study, our participants slowly whispered the syllables and stretched them before giving an answer. This may have made it easier to discover the phonemes in a VC than a CV. In a VC, the articulatory gestures allow the speaker to insert a short silence before the final consonant without disrupting the normal sequence of articulatory movements. In contrast, inserting a pause in a "stretched" CV more strongly disrupts the pattern of continuous articulatory gestures, since it is unnatural to pronounce a consonant in isolation (see Geudens & Sandra, 2003, for further comments). This may be especially so in items with stops. Stops are characterized by a complete stricture in the vocal tract, followed by a short explosion of air being pushed out from the lungs (Laver, 1994; Rietveld & Van Heuven, 1997). In the CV this "stop" or occlusion takes place initially which may strengthen the coarticulation of the C and V after the occlusion (e.g., Sussman, Bessell, Dalston, & Majors, 1997). As a result, the segmentation between consonant and vowel may be hampered. In contrast, in a VC, the occlusion in a stop is situated just in between the vowel and consonant and hence may be perceived as a natural pause in the speech signal. Thus, an interpretation that includes articulatory-phonetic factors may not only account for the advantage of VC over CV segmentation but may also help us to explain the disadvantage of stops over fricatives and liquids in CVs (see also Ball & Blachman, 1991).

2.2.2. Different or similar development of children with good and poor VC and CV segmentation skills?

Given our finding that CV segmentation is more difficult than VC segmentation, we wondered whether children with poor segmentation skills would show the same developmental path from VCs to CVs as children with better segmentation skills. One possibility is that poor segmenters are going through the same path, but arrive later at the onset of explicit phonological awareness than better segmenters. Thus, they may merely show a developmental delay, implying that if they had had an equal start position, they would progress from VCs to CVs at the same pace as good segmenters. There could be more than one reason for such a developmental lag. Perhaps these children are dispositionally slower in their phonological development, or maybe they only had less contact with written material, making it harder for them to abstract from surface phonetic features. Alternatively, if poor segmenters are characterized by a *qualitatively different* development of segmentation skills, the transition from VC to CV segmentation poses comparatively increasing difficulties for these children. They might lack sufficient insight, at this point, to break through the "phonetic glue" that characterizes CVs more than VCs, which could suggest a lower level of phonological sensitivity (cf. Bowey, 1994).

In order to investigate both scenarios, we matched a poor segmentation group with a good segmentation group on their VC segmentation skill, and then compared their ability to segment CVs. The rationale behind our matching procedure is analogous to the frequently used Reading Level Match (RLM) design in dyslexia research, in which dyslexic readers are matched with younger normal readers on a measure of reading ability, and compared on a criterion processing task (e.g., non-word reading) (cf. Bryant & Goswami, 1986; Jackson & Butterfield, 1989; for reviews of the design). In our study, we matched poor segmenters' VC scores of September with good segmenters' VC scores of June. Given both groups' equal start position on VC segmentation, a comparison of their CV segmentation performance enables us to evaluate the two proposed scenarios. If a developmental delay is characteristic for the poor group, their CV scores should equal those of the good group. However, if the poor group's segmentation difficulties reflect a qualitatively different development, the scores for the CVs will still be significantly lower in the poor than in the good group, despite the matching of both groups on VC scores.

We defined our segmentation groups in different ways. In our first analysis, poor segmenters were defined as those children who obtained zero scores at VC and CV segmentation in June ($n = 14$). The group with ''good'' segmentation skills consisted of the remaining children ($n = 27$), who were able to perform at least one correct VC or CV segmentation in June. Only complete segmentations, i.e., two segments isolated, were counted as correct (Complete segmentation criterion 1). We also performed a second analysis, in which the good group ($n = 20$) was defined as those children scoring above the median (>7) of the combined VC and CV scores in June. The poor segmentation group was the same as in the first analysis (Complete segmentation criterion 2). A third analysis was based on a more ''fine-grained'' segmentation measure. This latter measure combined the children's complete segmentation scores with their single segment responses (i.e., only one segment correctly reported). In the fine-grained analysis, the poor group ($n = 12$) was selected as scoring between zero and five correctly isolated segments, and the good group ($n = 20$) was selected above the median score (>37). Table 2 presents the mean scores, standard deviations, and ranges of the groups using the three criteria.

In our matching analyses, we used a continuous regression-based procedure instead of the classic matching design. This regression method was proposed by Jackson and Butterfield (1989) and first used in a study by Stanovich and Siegel (1994). The method is preferred because it avoids some of the substantial problems associated with the classic matching design, such as the arbitrariness of the matching (see Van den Broeck & Van den Bos, 2004). Basically, in the first step of this hierarchical regression-analysis, the criterion variable (CV segmentation) is regressed on the control variable (VC segmentation). After partialling out the variance due to VC segmentation ability, a variable designating group membership (poor vs. good) is entered into the equation. A significant regression coefficient (β) associated with group membership reflects a specific CV segmentation ''problem'' of the poor group. A potential difficulty of the regression technique when comparing groups of poor and good segmentation skills is the existence of partially non-overlapping distributions. In that case, participants of one group are compared with statistically extrapolated but non-existent participants of another group, resulting in meaningless conclusions (cf. Foorman, Francis, Fletcher, & Lynn, 1996). For example, when no children with good segmentation skills are found in the lower tail of the distribution of VC scores, the children with poor segmentation skills falling in that region are compared with statistically predicted scores based on the performance of children in the good group with high VC scores. To avoid this problem, all statistical analyses are exclusively based on the overlapping part of the distributions.

In order to test for a difference in CV segmentation between the statistically VC-matched groups, we compared two R^2s. The first R^2 expresses the explained variance in CV scores by VC segmentation ability, whereas the second R^2 additionally includes the variance explained by group membership. The difference between the two R^2s indicates the increment in the proportion of explained variance due to group membership. Using Complete segmentation criterion 1, this increment is .041, $F(1, 38) = 4.80$, $p = .035$, which means that group membership explains about 4% of the variance in CV scores, after variability in VC segmentation has been

Table 2
Means (*M*), standard deviations (*SD*), and ranges for the group selection criteria

Group selection criterion	Groups								
	Total group			Poor group			Good group		
	M	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range
Criterion 1[a]	13.46	(13.33)	0–32	0	(0)		20.44	(11.19)	1–32
Criterion 2[a]	13.46	(13.33)	0–32	0	(0)		26.20	(5.90)	11–32
Fine-grained criterion[b]	33.97	(25.67)	0–64	.75	(1.76)	0–5	57.55	(6.76)	39–64

[a] Complete segmentation criterion, maximum score = 32, *Mdn* = 7.
[b] Fine-grained segmentation criterion, maximum score = 64, *Mdn* = 37.

accounted for. Given the criterion that good segmenters are children who scored at least one VC or CV item correct in June, it is difficult to maintain that all children of the good group were really skillful at segmentation. Therefore, we performed an additional regression analysis in which the children with good segmentation skills were selected above the median of the summed VC and CV scores (Complete segmentation criterion 2). Using the same regression-based method, we again obtained a significant effect of group membership, $F(1,26) = 4.72$, $p = .039$, explaining now 8.9% of extra variance. Finally, the analysis based on the fine-grained measure revealed a significant group effect as well, $F(1,24) = 6.00$, $p = .022$, explaining 12% of extra variance.

These results suggest that the developmental transition of VC into CV segmentation ability is different for the children with poor segmentation skills. Two possible interpretations of this qualitatively different development can be proposed. Both accounts are related to the two variables on which the good and poor groups were not equated, i.e., age, and segmentation ability relative to age. Firstly, the most straightforward interpretation is to consider the relative inability of the poor group to break up the phonetically more complex CVs as a direct indication of a dispositional weaker phonological awareness. This account presumes that the matching of VC scores actually represented a matching of the children's phonological ability to segment VCs. A more refined interpretation than the previous one takes account of the possibility that the level of the VC scores of the poor group in September is partly the consequence of the explicit reading instruction that started in September. Even though the good group's VC segmentation scores in June was matched to the poor group's VC segmentation scores in September, these scores may have come about in a different way. For children in the good group, the insight in the segmental structure of words may have developed before they entered primary school through experience with letters and environmental print at home (Johnston, 1998). Thanks to their phonological sensitivity they may have realized in a spontaneous way that some of these letter sounds corresponded to sounds in spoken words, or they may have received explicit instruction at home (Adams, 1990; Bowey, 1994). Children in the poor group may have had less opportunities at home to learn about letters and print and/or may have needed more explicit instruction to break up the continuous speech signal. Still, this account indicates a qualitative difference between the poor and good group. The observed difference between both groups on the CV scores implies that the amount of explicit instruction that was needed in the poor group to master VC segmentation was not sufficient yet to break up CVs. Hence, the poor segmenters' lower degree of sensitivity to explicit instruction in the case of CVs also supports their qualitatively different development of VC and CV segmentation and could also indicate a lower degree of phonological skill.

2.2.3. Individual differences in VC–CV segmentation in relation to early reading skills

Our final purpose was twofold. First, we focused on the entire range of segmenters, and studied how their individual differences in VC and CV segmentation were related to their early reading skills. Second, we specifically focused on the poor and good segmentation groups, and explored whether their different development from VC to CV segmentation was also reflected in early reading performance. Children who arrive later at the segmentation of VCs, and again require more learning opportunities to

Table 3
Correlations, necessity (NI) and sufficiency (SI) indices for the relationship with reading proficiency

Measures	Reading Time 1					Reading Time 2				
	r	NI	(*SE*)	SI	(*SE*)	*r*	NI	(*SE*)	SI	(*SE*)
Complete segmentation										
VC June	.46	.56	(.195)	−.66	(.245)	.26	.20	(.151)	−.46	(.171)
CV June	.34	.48	(.236)	−.68	(.255)	.16	.16	(.172)	−.63	(.200)
VC + CV June	.41	.53	(.198)	−.54	(.207)	.21	.23	(.158)	−.53	(.179)
VC Sep	.30	.72	(.208)	−.92	(.221)	.28	.26	(.343)	−.59	(.158)
CV Sep	.41	.65	(.148)	−.67	(.182)	.30	.37	(.192)	−.44	(.169)
VC + CV Sep	.38	.67	(.161)	−.68	(.216)	.31	.44	(.263)	−.44	(.168)
Fine-grained segmentation										
VC June	.48	.65	(.136)	−.55	(.186)	.27	.30	(.159)	−.47	(.167)
CV June	.42	.60	(.154)	−.40	(.178)	.24	.36	(.150)	−.43	(.184)
VC + CV June	.46	.69	(.155)	−.52	(.175)	.26	.33	(.141)	−.37	(.171)
VC Sep	.27	.78	(.174)	−.92	(.203)	.28	.72	(.278)	−.73	(.143)
CV Sep	.38	.78	(.137)	−.77	(.213)	.31	.63	(.223)	−.57	(.163)
VC + CV Sep	.34	.81	(.159)	−.84	(.212)	.30	.57	(.252)	−.60	(.151)
Letter sounds	.75	.77	(.138)	.37	(.106)	.62	.40	(.118)	.46	(.100)

Note. Sep, September; N, necessity; S, sufficiency. All correlations larger than .30 are significant at the .05 level. NI and SI are based on Van den Broeck's (2004) measures for continuous variables (see Appendix B).

discover the phonemes in CVs may show problems in mastering GPCs and early self-teaching decoding skills as well (cf. Ehri, 1998; Share, 1995).

Correlations between the scores on the VC and CV segmentation tasks in June and September and the reading tests at Time 1 (February) and Time 2 (June) are presented in Table 3 (see Section 2.1.3 for a detailed description). As indicated by the high intercorrelations between the VC and CV scores in June, $r(41) = .95$, and in September, $r(41) = .88$, both measures were consistently tapping a similar skill. Yet, because VCs were easier to master for the children than CVs, the correlations with the reading scores varied with the developmental level of the children in phoneme segmentation skill, and their exposure to reading instruction (cf. Bowey, 2002). For instance, at the children's developmental level in June, both fine-grained CV ($r = .42$) and VC scores ($r = .48$) correlated moderately strongly with the reading scores at Time 1 (ps $< .008$). Nevertheless, for many children, complete segmentation was too difficult, especially for CVs, which may explain the relatively lower correlation with this measure ($r = .34$, $p = .029$). Having received reading instruction in September, VC scores rose to ceiling level. At that time, complete CV segmentation scores ($r = .41$, $p = .008$) were more "strongly" related to early reading than fine-grained VC scores ($r = .27$, $p = .092$). Note that the correlations with reading performance at Time 2 were much lower, which is in line with Vandervelden and Siegel's (1995) finding that rudimentary levels of skill in explicit phonological awareness (cf. segmentation of two-phoneme syllables) correlate to rudimentary phases in early decoding and word-reading skill, but show a weaker relationship with later levels of word reading.

An important problem in this examination of the relationship with early reading performance is that correlation suggests a linear relationship between variables with similar distributions. Yet, in our data, the distribution of the reading variable and the segmentation variable differed considerably from each other. As a matter of fact, distributions of VC and CV segmentation scores were bimodal. This could not be due to an inappropriate measurement of segmentation—given very high reliability estimates of our task—but reflects the nature of the children's segmentation patterns. For instance, in June, 12 of the 41 children could isolate maximally five segments, whereas the remaining children obtained scores above 14, leaving a natural cut-off line in the scores' distribution (e.g., Fig. 1). Such non-normal distributions in the segmentation task pose problems for the interpretation of Pearson correlations. These are bounded by the similarity in the variables' distributions, and their significance test is built on the assumption that the variables are normally distributed. Stahl and Murray (1998) argued that this problem is often ignored: "Although most research in this area analyzes [variables such as segmentation] as if they have a full range and are normally distributed, these are not variables with a full range and their distribution tends to be skewed" (p. 73). Moreover, correlation exclusively reflects the extent of the linear relationship between the variables. Hence, correlation may be an inadequate measure of the contingency between VC and CV segmentation and early reading skills (cf. Gough, Larson, & Yopp, 2000). Looking at Fig. 1, for example, it seems that the ordinary correlation has seriously underestimated the contingency of almost perfect necessity: none of the children who had problems in the segmentation task obtained good reading scores at Time 1. At the same time, good performance on the segmentation task was no guarantee at all that a child obtained high reading levels, i.e., it was not sufficient.

In order to avoid these statistical problems with correlations, we analyzed our data in a different way by combining inspection of the scatterplots with measures that quantify the concepts of necessary and sufficient conditions.[6] To our knowledge, these measures have only been used before in a study by Gough et al. (2000). [Stahl and Murray (1998) proposed that these measures can be calculated by using McNemar's Test, but they did not report the values.] Gough et al. calculated measures for a binary case, by separating the value of the phi coefficient into phiN (for necessity) and phiS (for sufficiency). In answer to the problem that these phi coefficients can only be used for dichotomous variables (consisting of only two values each), Van den Broeck (2004) proposed two new indices that take into account the continuity of the variables. Mathematical formulas are presented in Appendix B. Table 3 shows the contingency coefficients based on Van den Broeck's measures of necessity and sufficiency for continuous variables. By looking at the scatterplots (Figs. 1 and 3) in combination with the indices, it becomes clear

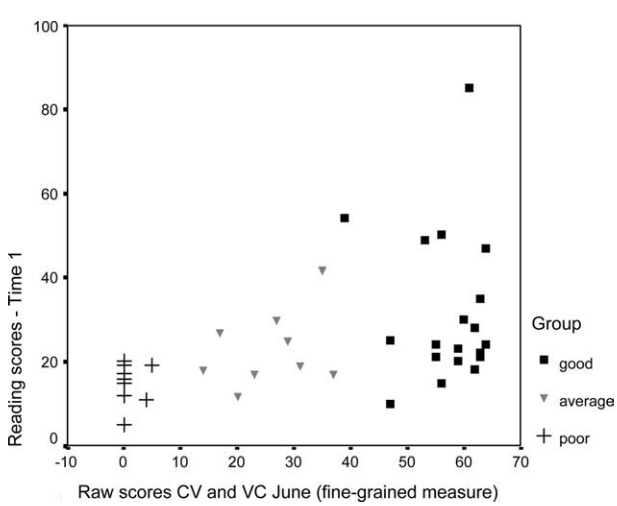


Fig. 1. Relationship CV–VC segmentation in June and later reading scores at Time 1.

[6] Importantly, a "necessary but not sufficient" relationship expresses contingency, and does not imply causality (see Section 3).

that the form of the relationship between VC and CV segmentation and early reading at Time 1 is much clearer than revealed in the correlations. There are children who cannot do either task very well and children who do both tasks well. There are also children who perform well on the segmentation and not on the reading test. But there are no children who perform well on the reading test and not on the segmentation test. This pattern gives rise to higher estimates of necessity in contrast with lower estimates of sufficiency. When related to the reading scores at Time 2 (Figs. 2 and 4), necessity is reduced, because there are poor segmenters who obtained moderate scores on the reading task. Such reduced pattern is not unanticipated. The reading test at Time 2 presumably measured more than children's rudimentary decoding skills, such as word-specific knowledge, experience with orthographic patterns and meaning (Adams, 1990; Ehri, 1998). In contrast, our segmentation task only measured children's rudimentary skill to break up syllables and isolate segments intentionally (see also Share, 1995; Vandervelden & Siegel, 1995).

In answer to our second question, we followed the reading performance of the poor and good segmentation groups separately. As all plots and analyses based on the three segmentation group criteria were highly similar, we will present the "fine-grained data" only. Student t tests confirmed that the different development of VC and CV segmentation for poor ($n = 12$) and good ($n = 20$) segmenters was reflected in their early reading performance. Poor segmenters obtained significantly lower scores on the reading test measured at Time 1, $t(30) = 3.02$, $p < .005$, and at Time 2 at the end of first grade, $t(30) = 1.65, p < .05$. Two patterns in the poor group are especially noteworthy. First, although children in the poor group strongly improved on the segmentation task in September (compare Figs. 1 and 3), their reading scores at Time 1 remained very low (Fig. 3). Second, even though these children managed to enhance their reading skills

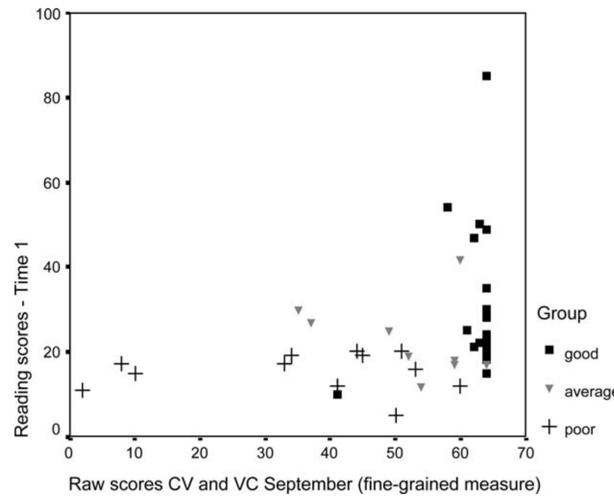


Fig. 3. Relationship CV–VC segmentation in September and later reading scores at Time 1.

from Time 1 to Time 2 onwards (e.g., compare Figs. 1 and 2, and 3 and 4), they did not seem to be capable to keep up with the good group. Their maximum reading scores in June concentrated around the mean, never reaching higher than half a standard deviation above the mean.

These findings also lend support to our interpretation of the matching analysis, which strongly emphasizes the interaction between children's two-phoneme segmentation development and "reading" instruction. Children in the good group were already able to segment simple two-phoneme syllables in June, suggesting that they had become aware of the segmental structure of words before the outset of reading instruction at school. This insight may have been acquired in a more spontaneous way (e.g., relating knowledge of some letter sounds to sounds in spoken words), or as a consequence of explicit instruction about letters and sounds at home. The fact that different experiences may underlie segmentation skill offers a possible account for the strong dispersion in the good

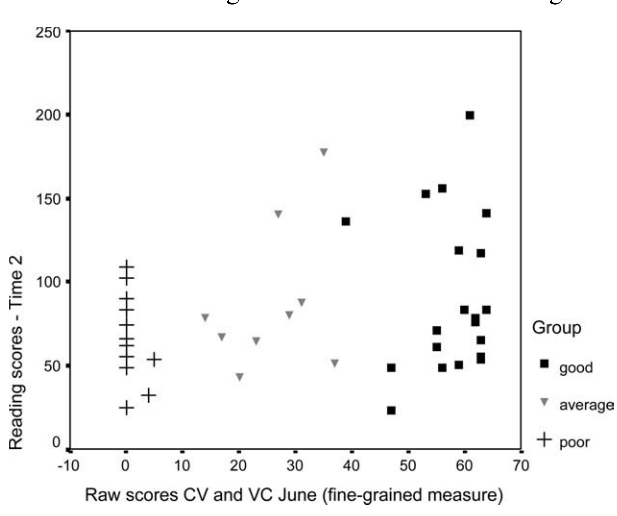


Fig. 2. Relationship CV–VC segmentation in June and later reading scores at Time 2.

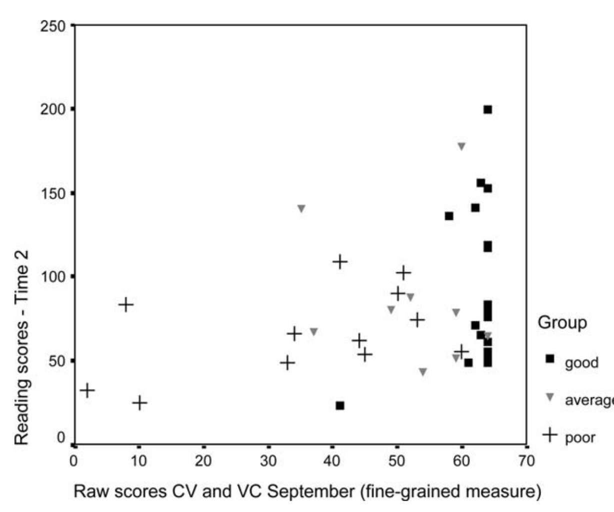


Fig. 4. Relationship CV–VC segmentation in September and later reading scores at Time 2.

segmentation group, reflecting a low sufficiency relationship between VC–CV segmentation and early reading (e.g., Fig. 1). The pattern in the poor group suggests that it is children's *spontaneous* insight into VC–CV segmentation that is indicative of good performance on reading measures. Children who only master VC and CV segmentation skill as a result of explicit instruction at school, remain in a lower range of reading scores (Figs. 3 and 4).

Finally, given the importance of these different experiences with print and GPCs, it is not surprising that letter sound knowledge, as measured at the outset of the first grade (see Section 2.1.3), is a strong predictor of children's early reading abilities, both measured at Time 1 ($r = .75$) and Time 2 ($r = .62$). As can be seen in Table 3, letter sound knowledge is more necessary for early reading skill than it is sufficient. Yet, it is more sufficient for early reading than VC–CV segmentation. Besides reflecting interests and informal experiences with print (Bowey, 1994), letter sound knowledge requires the skill of being able to associate sounds to an orthographic representation. This skill is central in early reading but not required in tasks that have been used to measure phonemic awareness. Other work has extensively supported that young children's letter knowledge is an excellent predictor of later reading and spelling (e.g., Adams, 1990; Caravolas et al., 2001; Muter, 1998).

3. General discussion

This study was concerned with individual differences in the early development of explicit phonemic awareness. To this end, we (1) followed children's performance to isolate segments in VCs and CVs, from the end of kindergarten (reported in Geudens & Sandra, 2003) to the start of first grade, (2) explored how segmentation development differed for poor and good segmenters, and (3) examined whether these differences were also reflected in the children's early reading ability.

Three findings are central in our research. First, we replicated our demonstration in Geudens and Sandra (2003) that Dutch speaking prereaders' early segmentation performance develops over the skill of VC segmentation to the skill of CV segmentation (cf. Bus, 1985; Uhry & Ehri, 1999; but see Schreuder & Van Bon, 1989). After having received phonics reading instruction for three weeks, children still found it harder to isolate the phonemes in CVs than in their reversed VCs (e.g., /to/ vs. /ot/). Importantly, we controlled for the vowel-like characteristics of consonants in initial and final position as much as possible, and paid attention to the children's perception of the spoken stimuli. The higher degree of complexity in CV vs. VC segmentation was explained in terms of perceptual- and articulatory-phonetic factors (see Geudens & Sandra, 2003). Such reasoning suggests that (1) phonetic factors rather than linguistic structures determine cohesion between phonemes, and that (2) the child's own articulation and perception assists in the gradual development of refined phonemic representations. These suggestions are in line with other demonstrations that phonetic factors play an important role in the development of phonological awareness (e.g., Caravolas & Bruck, 2000; Thomas & Sénéchal, 1998), early reading (e.g., Castiglioni-Spalten & Ehri, 2003; Rack, Hulme, Snowling, & Wightman, 1994), and spelling (e.g., Read, 1986).

A second important finding was related to individual differences in the VC–CV segmentation data. Children who had problems in breaking through the "phonetic glue" of two-phoneme syllables showed a different developmental path from VC to CV segmentation than children with better segmentation skills. Even when both groups were statistically matched on VC segmentation, their performance on the reversed CVs was still not equated: Poor segmenters found it harder to move from VC to CV segmentation. A closer inspection of the scatterplots suggested that this qualitatively different development might be influenced by experience-based differences. Although the matching technique equated poor segmenters' VC scores in September to good segmenters' VC scores in June, it did not equate the way in which these scores came about. Whereas good segmenters were already able to segment simple two-phoneme syllables before the outset of reading instruction at school, poor segmenters only reached the good group's VC segmentation level in the context of explicit phonics reading instruction. From this perspective, the results of the matching analyses indicate that poor segmenters needed explicit instruction to discover the "sounds" in VCs and that they needed comparatively more instructional effort to move to the more abstract skill of CV segmentation. Although the validity of this hypothesis should be confirmed with a larger group of participants, such an account may suggest that our poor segmenters had a weaker level of phonological sensitivity (see Bowey, 1994).

The third major finding in our study was that the lower segmentation skills in the poor vs. the good group had implications for the children's early reading performance. When related to the segmentation scores in June, children in the poor group performed significantly weaker on a reading test that was administered after six months of reading instruction. Remarkably, when their reading performance was related to the segmentation scores in September, it remained inferior, despite their notable improvement of segmentation skills. Even when they had the chance to enhance their reading skills near the end of first grade, they still did not seem to be able to exceed the average reading level. Although it is not clear yet whether this pattern extends to later stages of learning to read as well, it may offer a meaningful proposal for children's initial development. The pattern could suggest that it is not children's ability to segment VCs and CVs per se that is indicative of early reading success, but that what really matters is whether children acquired this skill more or less "*spontaneously*." This may have been the case for some of the good segmenters. "Children who are sensitive to the phonological structure of words may more readily understand the function of letters in print as representing sounds" (Bowey, 1994, p. 154), and hence develop segmentation skill in a more informal context of print-related experiences (see Adams, 1990, p. 84–85). These children may also have been ahead in the development of early decoding skills (Ehri, 1998; Share, 1995). Other children in the good group, however, may have obtained high segmentation scores due to intensive instruction about letters and sounds at home before the outset of formal reading instruction. Such tuition may have masked individual differences which are important for early reading development. This offers *one* account for the fact that high scores on the easy segmentation task in June did not automatically result in high scores on the reading tests (cf. Reitsma, 2002). Obviously, knowing how to segment VCs and CVs explicitly is not "sufficient" for early reading success (cf. Byrne, Fielding, Barnsley, & Ashley, 2000), which may also explain why the relationship with later reading measures was reduced. Reading also requires that children establish automatic, precise, and redundant connections between print and speech at fine-grained, larger subword, and word levels[7] (Ehri, 1998; Perfetti, 1992; Reitsma, 1997; Van den Broeck, 1997).

Despite their limitations, our findings add to the discussion of the relationship between measures of explicit phonological awareness and early reading abilities. Although the importance of this relation is well established, at least in studies with English children (see Adams, 1990; Share, 1995; for a review), it has not been explored before by following children's developmental differences in VC and CV segmentation skills. Furthermore, several studies with children learning to read in a relatively *transparent* orthography such as Dutch reported that measures of phonological awareness in kindergarten are not or only weakly related to individual differences in early reading skill (e.g., Braams & Bosman, 2000; De Jong & Van der Leij, 1999; Reitsma, 2002). In contrast, our findings show that a strong contingency between early segmentation abilities and later reading skills can also be revealed in this context, at least when tasks are used that are sensitive to the child's developmental level, and when measures are used that take into account the variables' range and distribution. This observed relationship was described as "necessary but not sufficient" and statistically supported by two indices for continuous variables (Van den Broeck, 2003). Several authors have confirmed the importance of this relationship (e.g., Gough et al., 2000; Smith, Simmons, & Kameenui, 2000; Stahl & Murray, 1998).

A cautionary note is in order concerning the interpretation of the concepts "necessary" and "sufficient." Whereas it may be tempting to look for a causal interpretation, these concepts merely describe a relationship of contingency in the data. Hence, when we use the term "necessary," we are not suggesting that children's insight in the segmental structure of words is an exclusive phonological skill that precedes the development of early "reading" skills. On the contrary, throughout the study, the children's segmentation development was interpreted in interaction with informal print-related experiences and explicit instruction about letters and sounds. What seemed to be "necessary" for the children's early decoding level was our *observed measure* of VC and CV segmentation, which developed "in tandem" with knowledge of letters and GPCs. What seemed to be most indicative for the children's reading success was a capacity to arrive at this knowledge and master segmentation skill in a more informal "literate" environment. In the absence of this insight, explicit reading instruction should offer those contexts that are most transparent to discover the relationship between letters and sounds so that precise mappings between graphemes and phonemes can be established.

In this respect, our study may also have implications for the design of beginning reading curricula. In the Netherlands and Flanders, phoneme segmentation is a central skill in most curricula for initial reading instruction (Mommers, 1990). In order to introduce new GPCs, children are taught to segment a visually presented word, mostly a CVC, in its corresponding segments, articulate the separate segments and then blend them again. Our results suggest that such training of correspondences between print and speech may be improved by introducing letters and sounds in a VC, starting with non-sonorous consonants (e.g., /ef/), and then introducing the same letters and sounds in the reversed CV (e.g., /fe/). Such instructions may especially be helpful for children with a lower level of phonological sensitivity as the children in our poor group. In the transparent phonetic VC context, these children may learn to discover the sounds, by stretching the syllable, articulating it slowly, and viewing the corresponding letters. Next, the teacher can introduce the reversed CV. Although, the sounds (phones) in the CV may be experienced as different from the sounds in the VC at first sight, the phonemes and graphemes are identical, which may make it easier for the poor segmenter to learn to abstract from phonetic factors and develop a detailed representation of the syllables' phonemes. Children may benefit from this insight in CVs to learn to segment more complex CVCs as well (Geudens & Sandra, 2003, Experiment 4; Lewkowicz & Low, 1979). Further research with a larger group of participants may shed light on the benefits of this technique for the initial teaching of GPCs and decoding.

[7] The possibility should also be mentioned that the poor group's segmentation skills in September were based on rather imprecise phonological representations, which may have hampered the children's initial self-teaching skill of corresponding orthographic patterns (Metsala & Walley, 1998; Share, 1995).

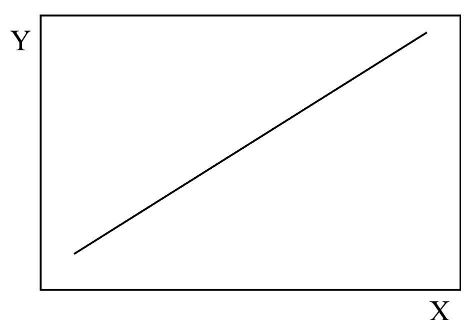
Appendix A. Experimental lists 1 and 2 with list number, stimulus, consonant type, and phoneme order

LIST 1				LIST 2			
Nr	Stimulus	Consonant Type	Phoneme Order	Nr	Stimulus	Consonant Type	Phoneme Order
21	/foː/	fricative	cv	18	/søː/	fricative	cv
27	/feː/ - *fee*	fricative	cv	36	/fyˑ/	fricative	cv
3	/suˑ/	fricative	cv	9	/saː/	fricative	cv
12	/fœy/	fricative	cv	26	/sɪˑ/ - *si*$^{+}$	fricative	cv
18	/øːs/	fricative	vc	21	/oːf/	fricative	vc
36	/yˑf/	fricative	vc	27	/eːf/ - *eef*	fricative	vc
9	/aːs/ - *aas*$^{+}$	fricative	vc	3	/uˑs/	fricative	vc
26	/iˑs/	fricative	vc	12	/œyf/	fricative	vc
1	/loː/	liquid	cv	6	/riˑ/	liquid	cv
32	/ryˑ/	liquid	cv	8	/lœy/ - *lui*	liquid	cv
19	/reː/ - *ree*	liquid	cv	40	/røː/ - *reu*$^{+}$	liquid	cv
24	/luˑ/	liquid	cv	14	/laː/ - *la*	liquid	cv
8	/œyl/ - *uil*	liquid	vc	1	/oːl/	liquid	vc
40	/øːr/	liquid	vc	32	/yːr/ - *uur*	liquid	vc
6	/iːr/	liquid	vc	19	/eːr/ - *eer*	liquid	vc
14	/aːl/ - *aal*$^{+}$	liquid	vc	24	/uˑl/	liquid	vc
10	/neː/ - *nee*	nasal	cv	5	/mœy/	nasal	cv
38	/noː/	nasal	cv	29	/myˑ/	nasal	cv
17	/miˑ/	nasal	cv	22	/naː/ - *na*	nasal	cv
31	/nuˑ/	nasal	cv	35	/nøː/	nasal	cv
5	/œym/	nasal	vc	10	/eːn/ - *een*	nasal	vc
29	/yˑm/	nasal	vc	38	/oːn/	nasal	vc
22	/aːn/ - *aan*	nasal	vc	17	/iˑm/	nasal	vc
35	/øːn/	nasal	vc	31	/uˑn/ - *oen*	nasal	vc
15	/puˑ/	plosive	cv	11	/toː/	plosive	cv
34	/teː/ - *thee*	plosive	cv	30	/paː/ - *pa*	plosive	cv
25	/pœy/	plosive	cv	2	/pøː/	plosive	cv
37	/tiˑ/	plosive	cv	20	/tyˑ/	plosive	cv
11	/oːt/	plosive	vc	15	/uˑp/	plosive	vc
30	/aːp/ - *aap*	plosive	vc	34	/eːt/ - *eet*	plosive	vc
2	/øːp/	plosive	vc	25	/œyp/	plosive	vc
20	/yˑt/	plosive	vc	37	/iˑt/	plosive	vc
7	/pɑk/ - *pak*	distractor	CVC	7	/kɑp/ - *kap*	distractor	CVC
16	/ɡyˑn/	distractor	CVC	16	/nyˑɡ/	distractor	CVC
39	/pɪl/ - *pil*	distractor	CVC	39	/lɪp/ - *lip*	distractor	CVC
4	/ɡɛɪk/	distractor	CVC	4	/kɛɪɡ/	distractor	CVC
13	/sɔm/ - *som*	distractor	CVC	13	/mɔs/ - *mos*	distractor	CVC
23	/bɛr/	distractor	CVC	23	/rɛb/	distractor	CVC
28	/kɪɡ/	distractor	CVC	28	/ɡɪk/	distractor	CVC
33	/lʏs/ - *lus*	distractor	CVC	33	/sʌl/ - *sul*	distractor	CVC

For items with word status, Dutch orthography is added in italics. Words with a "+" are unfamiliar to prereaders and beginning readers (Schaerlaekens, Kohnstamm, & Lejaegere, 1999).

Appendix B

The Necessity Index (NI) gives an indication of the extent that a continuous variable X is a necessary condition for a continuous variable Y. Importantly, such relationship expresses contingency, and does not imply causality. Being a necessary condition for Y means that in order to score high on Y it is necessary to score high on X. Thus, scores low on X and high on Y violate this requirement. However, scores high on X and low on Y are not inconsistent with the necessity requirement, because the possibility exists that good scores on other variables than X are required to raise the Y-score. Analogously, the Sufficiency Index (SI) reflects the extent that X is a sufficient condition for Y. To meet the sufficiency condition implies that high X-scores are always associated with high Y-scores, but low X-scores may be associated with high Y-scores, meaning that X is not a necessary condition for Y. Our interpretation of the necessity and sufficiency condition is a linear continuous one. For example, to meet the necessity requirement, a medium score on Y requires at least a medium score on X. The diagonal in the figure, which connects the points constructed by the extreme X and Y values, depicts this linear interpretation: The necessity condition requires that no data points are situated above the diagonal, and the sufficiency condition demands that no points are lying beneath the diagonal.



Accordingly, a strong correlation implies high necessity and high sufficiency relationships. NI and SI are mathematical expressions of the extent that the necessity or sufficiency requirement is fulfilled. They are calculated by averaging over a weighted sum of ratios. Each ratio expresses the distance of each point to the diagonal relative to the maximum possible distance to the diagonal, i.e., the difference between the maximum value on Y (NI) or the minimum value on Y (SI) and the Y-value on the diagonal. Both indices vary theoretically between −2 and 1, although in almost all cases they will vary between −1 and 1. An index of zero, which corresponds to the situation where all data points are scattered homogeneously over the entire space of the figure, indicates the absence of a necessity (NI) or sufficiency (SI) relationship. Negative values of NI or SI are found when even more data points violate the necessity or sufficiency requirement, and thus are generally situated closer to the maximum (for NI) or minimum value of Y (for SI). When both indices are negative, they are generally associated with a negative correlation. Further, a comparison of both indices is useful to formulate combined conclusions, e.g., "necessary but not sufficient condition," or vice versa. Finally, it is of major importance to ascertain that extreme data points are not unreliable outliers because this technique is very sensitive to the position of extreme data points.

When the values of X vary between the minimum value m_1 and the maximum value M_1, and the values of Y vary between the minimum m_2 and the maximum M_2, and when the data points (x_i, y_i) fulfill the following condition (n_1 points), that is, they are situated above the diagonal,

$$\frac{[(M_2 - m_2)x_i + M_1m_2 - M_2m_1]}{M_1 - m_1} - y_i < 0$$

then, the distance between each point and the diagonal is

$$\Delta_{(y_i - D_i)} = [y_i(M_1 - m_1) - (M_2 - m_2)x_i - M_1m_2 + M_2m_1]/(M_1 - m_1)$$

and the distance between the maximum M_2 and the diagonal is

$$\Delta_{(M_2 - D_i)} = [M_2(M_1 - m_1) - (M_2 - m_2)x_i - M_1m_2 + M_2m_1]/(M_1 - m_1)$$

$$\text{NI} = 1 - \frac{3}{n_1}\sum_{i=1}^{n_1}\frac{(M_1 - x_i)\Delta_{(y_i - D_i)}}{(M_1 - m_1)\Delta_{(M_2 - D_i)}}.$$

When no data points fulfill the condition, NI = 1.

When the data points (x_i, y_i) meet the following condition (n_2 points), that is, they are situated beneath the diagonal

$$\frac{[(M_2 - m_2)x_i + M_1m_2 - M_2m_1]}{M_1 - m_1} - y_i > 0$$

then, the distance between each point and the diagonal is

$$\Delta_{(D_i - y_i)} = [y_i(m_1 - M_1) + (M_2 - m_2)x_i + M_1m_2 - M_2m_1]/(M_1 - m_1)$$

and the distance between the minimum m_2 and the diagonal is

$$\Delta_{(D_i - m_2)} = [m_2(m_1 - M_1) + (M_2 - m_2)x_i + M_1m_2 - M_2m_1]/(M_1 - m_1),$$

$$\text{SI} = 1 - \frac{3}{n_2}\sum_{i=1}^{n_2}\frac{(x_i - M_1)\Delta_{(D_i - y_i)}}{(M_1 - m_1)\Delta_{(D_i - m_2)}}.$$

Again, when no points fulfill the condition, SI = 1.

References

Adams, M. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Ball, E. W., & Blachman, B. A. (1991). Does phoneme awareness training in kindergartners make a difference in early word recognition and developmental spelling? *Reading Research Quarterly, 26*, 49–66.

Blevins, J. (1996). The syllable in phonological theory. In J. A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 206–244). Cambridge, MA: Blackwell.

Bowey, J. A. (1994). Phonological sensitivity in novice readers and nonreaders. *Journal of Experimental Child Psychology, 58*, 134–159.

Bowey, J. A. (2002). Reflections on onset-rime and phoneme sensitivity as predictors of beginning word reading [Electronic version]. *Journal of Experimental Child Psychology, 82*, 29–40.

Braams, T., & Bosman, A. M. T. (2000). Fonologische vaardigheden, geletterheid en lees- en spellinginstructie [Phonological skills, literacy, reading and spelling instruction]. *Tijdschrift voor Orthopedagogiek, 39*, 199–211.

Bryant, P. E., & Goswami, U. (1986). Strengths and weaknesses of the reading level design: A comment on Backman, Mamen, and Ferguson. *Psychological Bulletin, 100*, 101–103.

Bus, A. G. (1985). Preparatory reading instruction in kindergarten. Some comparative research into methods of auditory and auditory–visual training of phonemic analysis and blending. *Pedagogische Studiën, 62*, 269–279.

Byrne, B., & Fielding-Barnsley, R. (1991). Evaluation of a program to teach phonemic awareness to young children. *Journal of Educational Psychology, 82*, 451–455.

Byrne, B., Fielding, Barnsley, R., & Ashley, L. (2000). Effects of preschool phoneme identity training after six years: Outcome level distinguished from rate of response. *Journal of Educational Psychology, 92*, 659–667.

Caravolas, M., & Bruck, M. (2000). Vowel categorization skill and its relationship to early literacy skills among first-grade Québec-French children. *Journal of Experimental Child Psychology, 76*, 190–221.

Caravolas, M., Hulme, C., & Snowling, M. J. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of Memory and Language, 45*, 751–774.

Castiglioni-Spalten, M. L., & Ehri, L. C. (2003). Phonemic awareness instruction: Contribution of articulatory segmentation to novice beginners' reading and spelling. *Scientific Studies of Reading, 7*, 25–52.

Coleman, E. B. (1970). Collecting a data base for a reading technology. *Journal of Educational Psychology, 61*, 1–63.

Collins, B., & Mees, I. (1981). *The sounds of English and Dutch*. The Hague, The Netherlands: Leiden University Press.

Content, A., Kolinsky, R., Morais, J., & Bertelson, P. (1986). Phonetic segmentation in prereaders: Effect of connective information. *Journal of Experimental Child Psychology, 42*, 49–72.

De Baar, K. (1995). *De Leeslijn—De leesweg* [Reading line—The reading path]. Amsterdam: Meulenhof Educatief.

De Jong, P. F., & Van der Leij, A. (1999). Specific contributions of phonological abilities to early reading acquisition: Results from a Dutch latent variable longitudinal study. *Journal of Educational Psychology, 91*, 450–476.

Duighuisen, H. C. M., Van Bon, W. H. J., & Schreuder, R. (1991). *Problematische fonemische segmentatie* [Problematic phonemic segmentation]. Eindrapport SVO-project 7012. Nijmegen, The Netherlands: Instituut voor Orthopedagogiek KUN.

Ehri, L. C. (1998). Grapheme–phoneme knowledge is essential for learning to read words in English. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 3–40). Mahwah, NJ: Erlbaum.

Foorman, B. R., Francis, D. J., Fletcher, J. M., & Lynn, A. (1996). Relation of phonological and orthographic processing to early reading: Comparing two approaches to regression-based, reading-level-match designs. *Journal of Educational Psychology, 88*, 639–652.

Fudge, E. C. (1987). Branching structure within the syllable. *Journal of Linguistics, 23*, 359–377.

Geudens, A., & Sandra, D. (2003). Beyond children's implicit phonological knowledge: No support for an onset-rime structure in children's explicit phonological awareness. *Journal of Memory and Language, 49*, 157–182.

Goldsmith, J. A. (1990). *Autosegmental and metrical phonology*. Oxford, England: Blackwell.

Gombert, J. E. (1996). What do children do when they fail to count phonemes? *International Journal of Behavioral Development, 19*, 757–772.

Gough, P. B., Larson, K. C., & Yopp, H. (2000). The structure of phonemic awareness. Retrieved February 22, 2000, from University of Texas, Department of Psychology. Available http://www.psy.utexas.edu/psy/klarson/recife.html.

Hatcher, P. J., Hulme, C., & Ellis, A. W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills. The phonological linkage hypothesis. *Child Development, 65*, 41–57.

Hohn, W. E., & Ehri, L. C. (1983). Do alphabet letters help prereaders acquire phonemic segmentation skill? *Journal of Educational Psychology, 75*, 752–762.

Hulsen, G., & Voets, C., 2003. De rol van foneemsynthese, letterverklanking, orthografische kennis, en gevoeligheid voor bigramfrequenties in de leesontwikkeling [The role of phoneme blending, letter-to-sound mapping, orthographic knowledge, and sensitivity to bigramfrequencies in reading development]. Unpublished Master's thesis, University of Leiden, The Netherlands.

Jackson, N. E., & Butterfield, E. C. (1989). Reading-level-matched designs: Myths and realities. *Journal of Reading Behavior, 21*, 387–412.

Johnston, R. S. (1998). The role of letter learning in developing phonemic awareness skills in preschool children: Implications for explanations of reading disorders. Development and disorders. In C. Hulme & R. M. Joshi (Eds.), *Reading and spelling* (pp. 287–301). Mahwah, NJ: Erlbaum.

Johnston, R. S., Anderson, M., & Holligan, C. (1996). Knowledge of the alphabet and explicit awareness of phonemes: The nature of the relationship. *Reading and Writing, 8*, 217–234.

Kager, R. (1989). *A metrical theory of stress and destressing in English and Dutch*. Dordrecht: Foris Publications.

Krashen, S. (2001). Does "pure" phonemic awareness training affect reading comprehension? *Perceptual and Motor Skills, 93*, 356–358.

Landerl, K., & Wimmer, H. (2000). Deficits in phoneme segmentation are not the core problem of dyslexia: Evidence from German and English children. *Applied Psycholinguistics, 21*, 243–262.

Laver, J. (1994). *Principles of phonetics*. Cambridge, UK: Cambridge University Press.

Lewkowicz, N., & Low, L. (1979). Effects of visual aids and word structure on phonemic segmentation. *Contemporary Educational Psychology, 4*, 238–252.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*, 431–461.

Martensen, H., Maris, E., & Dijkstra, T. (2000). When does inconsistency hurt. On the relation between phonological-consistency effects and the reliability of sublexical units. *Memory & Cognition, 28*, 648–656.

Martinot, C., & Gombert, J. E. (1996). Le développement et le contrôle des connaissances phonologiques'al'âge préscolaire [The development and control of pre-readers' phonological knowledge]. *Revue de Neuropsychologie, 6*, 251–270.

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Mahwah, NJ: Erlbaum.

Mommers, M. J. C. (1990). Metalinguistic awareness and learning to read. In P. Reitsma & L. Verhoeven (Eds.), *Acquisition of reading in Dutch* (pp. 29–42). Dordrecht: Foris.

Morais, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy training and speech segmentation. *Cognition, 24*, 45–64.

Morais, J., Mousty, P., & Kolinsky, R. (1998). Why and how phoneme awareness helps learning to read. Development and disorders. In C. Hulme & R. M. Joshi (Eds.), *Reading and spelling. Development and disorders* (pp. 127–152). Mahwah, NJ: Erlbaum.

Muter, V. (1998). Phonological awareness: Its nature and its influence over early literacy development. In C. Hulme & R. M. Joshi (Eds.), *Reading and spelling. Development and disorders* (pp. 113–125). Mahwah, NJ: Erlbaum.

Nation, K., & Hulme, C. (1997). Phonemic segmentation, not onset-rime segmentation, predicts early reading and spelling skills. *Reading Research Quarterly, 32*, 154–167.

Patel, T. K. (2001). Individual differences in learning to read in English and Dutch children. Unpublished doctoral dissertation. University of York, England.

Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 145–174). Hillsdale, NJ: Erlbaum.

Perfetti, C. A., Beck, I., Bell, L. C., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. *Merrill-Palmer Quarterly, 33*, 283–319.

Pierrehumbert, J., & Nair, R. (1995). Word games and syllable structure. *Language and Speech, 38*, 77–114.

Rack, J. P., Hulme, C., Snowling, M. J., & Wightman, J. (1994). The role of phonology in young children learning to read words: The direct mapping hypothesis. *Journal of Experimental Child Psychology, 57*, 42–71.

Read, C. (1986). *Children's creative spelling*. London: Routledge.

Reitsma, P. (1997). How to get friends in beginning word recognition. In C. K. Leong & R. M. Joshi (Eds.), *Cross-language studies of learning to read and spell* (pp. 213–233). Dordrecht: Kluwer.

Reitsma, P. (2002). Precursors of phonemic awareness. In L. Verhoeven, P. Reitsma, & C. Elbro (Eds.), *Precursors of functional literacy* (pp. 33–48). Amsterdam: John Benjamins.

Rietveld, A. C. M., & Van Heuven, V. J. (1997). *Algemene fonetiek* [General phonetics]. Bussum, The Netherlands: Coutinho.

Schaerlaekens, A. M., Kohnstamm, D., & Lejaegere, M. (1999). *Streeflijst woordenschat voor zesjarigen. Derder herziene versie gebaseerd op nieuw onderzoek in Nederland en België* [Target word list for 6-year-olds. Third revision based on new research in the Netherlands and Belgium]. Lisse: Swets & Zeitlinger.

Scholes, R. J. (1998). The case against phonemic awareness. *Journal of Research in Reading, 21*, 177–189.

Schreuder, R., & Van Bon, W. H. J. (1989). Phonemic analysis: Effects of word properties. *Journal of Research in Reading, 12*, 59–78.

Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition, 55*, 151–218.

Skjelfjord, V. J. (1987). Phonemic segmentation. An important subskill in learning to read: I. *Scandinavian Journal of Educational Research, 31*, 41–58.

Smith, S. B., Simmons, D. C., & Kameenui, E. J. (2000). Synthesis of research on phonological awareness: Principles and implications for reading acquisition. Retrieved February, 21, 2000 from University of Oregon. Available http://idea.uoregon.edu/~ncite/documents.techrep.tech21.html.

Sproat, R., & Fujimura, O. (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics, 21*, 291–311.

Stahl, S. A., & Murray, B. A. (1998). Issues involved in defining phonological awareness and its relation to early reading. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 65–89). Mahwah, NJ: Erlbaum.

Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology, 86*, 24–53.

Sussman, H. M., Bessell, N., Dalston, E., & Majors, T. (1997). An investigation of stop place of articulation as a function of syllable position: A locus equation perspective. *Journal of the Acoustical Society of America, 101*, 2826–2838.

Thomas, E. M., & Sénéchal, M. (1998). Articulation and phoneme awareness of 3-year-old children. *Applied Psycholinguistics, 19*, 363–391.

Treiman, R. (1992). The role of intrasyllabic units in learning to read and spell. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 65–106). Hillsdale, NJ: Lawrence Erlbaum.

Trommelen, M. (1984). *The syllable in Dutch. With special reference to diminutive formation*. Dordrecht: Foris.

Uhry, J. K., & Ehri, L. C. (1999). Ease of segmenting two- and three-phoneme words in kindergarten: Rime cohesion or vowel salience? *Journal of Educational Psychology, 91*, 594–603.

Van Bon, W. H. J., & Duighuisen, H. C. M. (1995). Sometimes spelling is easier than phonemic segmentation. *Scandinavian Journal of Psychology, 36*, 82–94.

Van den Broeck, W. (1997). De rol van fonologische verwerking bij het automatiseren van de leesvaardigheid [The role of phonological processing in the acquisition of fluent and automatic reading]. Unpublished doctoral dissertation, University of Leiden, The Netherlands.

Van den Broeck, W. (2003). A necessity index and a sufficiency index for the description of the contingency between two continuous variables (manuscript in preparation).

Van den Broeck, W., & Van den Bos, K. P. (2004). The nonword-reading deficit of disabled readers: Fact or artifact? (Submitted for publication).

Vandervelden, M. C., & Siegel, L. S. (1995). Phonological recoding and phoneme awareness in early literacy: A developmental approach. *Reading Research Quarterly, 30*, 854–875.

Van der Wissel, A. (1984). Foneemsegmentatie als kernprobleem van ernstig falende GLO-leerlingen [Phoneme segmentation as core problem in children with learning difficulties]. *Tijdschrift voor Orthopedagogiek, 23*, 247–257.

Van Reenen, P., & Jongkind, A. (2000). De vocalisering van de /l/ in het Standaard- Nederlands [Vocalization of /l/ in Standard Dutch]. *Taal en Tongval, 52*, 189–199.

Verhoeven, L. (1993). *Cito Drie-minuten-toets. Handleiding* [Cito Three-minute-test. Manual]. Arnhem: Cito.

Yavas, M. S., & Gogate, L. J. (1999). Phoneme awareness in children: A function of sonority. *Journal of Psycholinguistic Research, 28*, 245–260.

Yip, M. (2002). Non-arguments for sub-syllabic constituents. *Journal of Chinese Phonology [Special Issue], 11*, 49–72.