Word Recognition and Nonword Repetition in Children with Language Disorders: The Effects of Neighborhood Density, Lexical Frequency, and Phonotactic Probability

Rispens, J.; Baker, A.; Duinmeijer, I.

DOI
10.1044/2014_JSLHR-L-12-0393

Publication date
2015

Document Version
Final published version

Published in
Journal of Speech, Language, and Hearing Research

JSLHR

Research Article

Word Recognition and Nonword Repetition in Children With Language Disorders: The Effects of Neighborhood Density, Lexical Frequency, and Phonotactic Probability

Judith Rispens,[a] Anne Baker,[a] and Iris Duinmeijer[a]

Purpose: The effects of neighborhood density (ND) and lexical frequency on word recognition and the effects of phonotactic probability (PP) on nonword repetition (NWR) were examined to gain insight into processing at the lexical and sublexical levels in typically developing (TD) children and children with developmental language problems.
Method: Tasks measuring NWR and word recognition were administered to 5 groups of children: 2 groups of TD children (5 and 8 years old), children with specific language impairment (SLI), children with reading impairment (RI), and children with SLI + RI (all 7–8 years old).
Results: High ND had a negative effect on word recognition in the older TD children and in the children with RI only. There was no ND effect in the younger children or in the children with SLI, who all had lower receptive vocabulary scores than the age-matched TD children and the RI groups. For all groups, NWR items with low PP were more difficult to repeat than items with high PP. This effect was especially pronounced in children with RI.
Conclusion: Both the stage of vocabulary development and the type of language impairment (SLI or RI) impact the way ND and PP affect word recognition and NWR.

Children with specific language impairment (SLI) often experience difficulties with lexical and phonological processes in the context of broader language impairments in language development. These difficulties surface, for instance, in tasks that tap word recognition and nonword repetition (NWR; e.g., Bishop, 1997; Criddle & Durkin, 2001; Edwards & Lahey, 1998; Gathercole, 2006; Maillart, Schelstraete, & Hupet, 2004; Mainela-Arnold, Evans, & Coady, 2008). Children with reading impairment (RI) have severe difficulties with the acquisition and automatization of literacy skills (Snowling, 2000). Investigations into the oral language abilities of children with RI have also indicated difficulties with word recognition (Metsala, 1997) and NWR (for a meta-analysis of NWR in RI, see Melby-Lervag & Lervag, 2012). SLI and RI are separate disorders but are related, as comorbid SLI and RI have been documented (Bishop & Snowling, 2004; Catts, Adlof, Hogan, & Ellis Weismer, 2005; McArthur, Hogben, Edwards, Heath, & Mengler, 2000). There is no consensus in the literature about the nature of the overlap between SLI and RI. Even though previous results have demonstrated difficulties in both groups of children with word recognition and NWR, it is not clear what this overlap means in terms of underlying linguistic abilities, as tasks measuring word recognition and NWR involve several linguistic skills. In the current study we examined three linguistic variables. We investigated the effects of neighborhood density and lexical frequency on word recognition and the influence of phonotactic probability on NWR in children with SLI and/or RI.

Both neighborhood density and lexical frequency affect word recognition (Luce & Pisoni, 1998). *Neighborhood density* (ND) is defined as the number of words differing in one phoneme from another word in terms of addition, substitution, or deletion in any word position (e.g., Luce & Pisoni, 1998). It has been demonstrated that phonotactic probability affects NWR (e.g., Munson, 2001). *Phonotactic probability* (PP) refers to the frequency of occurrence of individual phonemes and phoneme combinations (biphones) within a given language (e.g., Storkel, Armbruster, & Hogan, 2006; Vitevitch & Luce, 1999). The effects of such variables give insight into the organization of the mental lexicon and the way information at the lexical and sublexical levels is processed. The

[a]Amsterdam Center for Language and Communication, University of Amsterdam, the Netherlands

Correspondence to Judith Rispens: J.E.Rispens@uva.nl

Editor: Rhea Paul
Associate Editor: Lizbeth Finestack

Received December 11, 2012
Revision received May 30, 2013
Accepted October 10, 2014
DOI: 10.1044/2014_JSLHR-L-12-0393

Disclosure: The authors have declared that no competing interests existed at the time of publication.

question remains whether the difficulties with word recognition (lexical level) and NWR (sublexical level) that have been observed in children with SLI and in children with RI are of the same nature and are the same in severity. The current study addresses this question by examining the influence of ND and lexical frequency on word recognition and the influence of PP on NWR in children with SLI, in children with RI, and in children with SLI and RI to evaluate similarities and dissimilarities in the nature of lexical and sublexical processing in these children.

Effects of ND and Lexical Frequency on Word Recognition in Typical Development

Recognizing words with many neighbors is slower compared with recognizing words that have few neighbors (e.g., Garlock, Walley, & Metsala, 2001; Goh, Suárez, Yap, & Tan, 2009). During auditory recognition, the target word needs to be discriminated from other lexical representations. Activation of phonological neighbors slows down the recognition of the target due to the competition between the activated lexical items. The competition effect is stronger for target words with many neighbors relative to targets with few neighbors. Upon processing these latter words, fewer related words are activated and therefore the competition effect is weaker. The disadvantage in recognition of words high in ND compared with words low in ND has been demonstrated in adults (Goh et al., 2009) using an auditory lexical decision task and in children of around 7 to 8 years old (Garlock et al., 2001) using a gating task, but it was not found in kindergarten children who performed a gating task (Garlock et al., 2001). Lexicons of children around kindergarten age are developing and are smaller than those of older children and adults. A strong competition effect is thus unlikely because there are relatively few neighbors. Over time, the influence of ND will increase as more vocabulary items are added to the lexicon, so that phonological neighborhoods become more densely populated and there will be more competition from phonological neighbors during word recognition. The effect of ND on word recognition is predicted to be more influential in larger vocabularies, not only because of the greater number of lexical items but also because an increase in vocabulary size leads to more phonologically detailed and robust lexical representations (see Metsala, 1999).

In addition to ND, lexical frequency influences the processing of lexical information. Words with high lexical frequency are processed faster and more accurately than words with low lexical frequency in adults and children (e.g., Luce & Pisoni, 1998; Walley, Michela, & Wood, 1995). Luce and Pisoni (1998) further observed that, in an experimental study with adults, effects of ND and lexical frequency interacted: High-frequency words with low ND were recognized at earlier points in time than high-frequency words in dense neighborhoods. The advantage reflects the facilitative effect of high frequency and the least competition of phonological neighbors. Mainela-Arnold et al. (2008) also observed this interaction between ND and frequency, in addition to a significant main effect of frequency, in children (ages between 8 and 12 years, mean age 10 years) who performed a gating task. In contrast, Garlock et al. (2001) did not find an interaction between ND and lexical frequency in a gating task (participants were 5 years old, 7 years old, and adults). In addition to lexical frequency, Garlock et al. studied the effects of the age of acquisition of lexical items on word recognition. They found that early acquisition of an item facilitated word recognition of that item. This effect may have masked a potential lexical frequency effect in their study. In the current study, we further investigated the effect of ND and lexical frequency on word recognition using a lexical decision task while we controlled for age of acquisition. The same cross-sectional developmental perspective as Garlock et al. was taken to investigate whether the effect of ND increases over time.

Effects of ND and Lexical Frequency on Word Recognition in Children With SLI and RI

As discussed above, Mainela-Arnold et al. (2008) found that, in a gating task, participants required less phonetic information to correctly recognize words high in lexical frequency compared with words low in lexical frequency. In the case of high-frequency words, words from sparse neighborhoods were identified at earlier gates than were words from dense neighborhoods. Children with SLI (mean age 10 years) also participated in that experiment, and Mainela-Arnold et al. found that the influence of ND and lexical frequency on word recognition was similar between the SLI and typically developing (TD) groups. Sensitivity to lexical frequency was also found in another study during word recall (tested with a sentence span task), where children with SLI recalled more high-frequency words than low-frequency words, just like the TD children (Mainela-Arnold, Evans, & Coady, 2010).

Mixed evidence has been observed regarding the effect of ND on word recognition in children with reading difficulties. Metsala (1997) found that, in a gating task, children with RI (around 9 years of age) needed more input to recognize words in sparse neighborhoods than did children without RI. Metsala concluded that spoken-word recognition in children with RI is affected because of less detailed phonological representations. In contrast, Griffiths and Snowling (2001), using an auditory gating task, did not find that children with RI were affected differently by ND, and they concluded that lexical processing is unaffected in children with RI. The authors noted that differences in sampling (e.g., age and vocabulary levels of the children) may have caused the conflicting results. Furthermore, Griffiths and Snowling speculated that word recognition is unproblematic for children with RI but that deficits are likely in children with a reading disability who have concomitant language impairment. Both the study of Metsala (1997) and Griffiths and Snowling (2001) found facilitating effects of high lexical frequency that were similar between the children with RI and the TD children.

In the present study, we investigated ND and lexical frequency in 7- and 8-year-old children with SLI and/or RI to obtain additional evidence on word recognition in children with SLI and RI. Our study adds to the current findings in two ways. First, the sample of children with language impairments in the current study was, on average, 2 years younger than the sample reported on by Mainela-Arnold et al. (2008). Our results give insight into lexical processing in children with SLI who are at an earlier developmental stage and thus have less developed lexicons. Second, in addition to a group of children with RI, we assessed children who experience RI as well as oral language impairment (SLI + RI) to investigate whether the presence of oral language impairment affects the influence of ND on word recognition, as proposed by Griffiths and Snowling (2001).

Effects of PP on the Repetition of Nonwords in Typical Development

Nonwords consisting of phonemes and sequences of phonemes that occur frequently in a language are easier to repeat than nonwords with low PP (Edwards, Beckman, & Munson, 2004; Munson, 2001; Thorn & Frankish, 2005; Vitevitch & Luce, 1998). Edwards et al. (2004) investigated the effect of PP on NWR in a developmental framework and compared children ages 3–8 years with adults. They found that the negative effect of low PP on the repetition of the nonwords was largest for children ages 3–4 years. Furthermore, NWR performance significantly correlated with vocabulary. This finding supports the view that lexical development drives the development of phonological representations, which become more finely grained over time (Metsala, 1999). These robust phonological representations contribute not only to NWR in general but especially to low PP segments. From detailed and robust phonological representations, generalizations can be made to aid repetition of infrequently occurring segments (Edwards et al., 2004). The current study further investigated the changing effect of PP on NWR over time and examined whether the effect of low PP is larger in children ages 5 and 6 years compared with children ages 7–8 years and whether this is associated with vocabulary size.

Effects of PP on the Repetition of Nonwords in SLI and RI

Low PP affects the repetition of nonwords significantly more in children with SLI than in TD children (Munson, Kurtz, & Windsor, 2005). These results suggest that children with SLI are sensitive to the frequency of phonotactic patterns in their language but that, relative to TD children, they have more problems with patterns occurring infrequently. An association between PP and vocabulary size was found in the group of children with SLI (Munson et al., 2005), which suggests that the larger effect of low PP may be related to their more limited vocabulary and the suggested weaknesses in their phonological representations (see also Coady, Mainela-Arnold, & Evans, 2013; McKean, Letts, & Howard, 2013). Coady, Evans, and Kluender (2010) found a main effect of PP on NWR but no interaction with group (children with SLI compared with TD children). Inconsistencies in findings between studies may be due to differences in sampling (e.g., age) but also to differences between reading ability of children with SLI. Data from recent studies suggest that NWR performance in children with SLI depends on the level of their literacy skills, considering that children with SLI and concomitant RI have been found to experience severe difficulties with repeating nonwords, in contrast to children with SLI but without RI, who performed similar to age-matched TD children (Bishop, Hayiou-Thomas, McDonald, & Bird, 2009; Catts et al., 2005; Rispens & Baker, 2012; Rispens & Parigger, 2010). Bishop et al. (2009) showed that children with SLI but without RI outperformed children with both SLI and RI at 9 years of age but not at 4 years of age (data analyzed in retrospect). The authors attribute this to differences in orthographic knowledge. They maintain that the acquisition of orthography provides feedback about the phonological system, thereby increasing phonological awareness and strengthening the phonological form of lexical representations (see also Goswami, 2000). As discussed above, NWR is positively influenced by a flexible and robust phonological system (Coady et al., 2013; Munson et al., 2005; Rispens & Baker, 2012). Thus, it is possible that variations in phonological abilities (related to differences in literacy abilities) cause variations in NWR ability across studies in children with SLI. It is not clear, though, whether the extent of the influence of PP, rather than the overall NWR accuracy, also depends on the reading abilities of children with SLI and whether low PP also affects children with RI. This was addressed in the current study by dividing the group of children with SLI into two groups: children with SLI only, whose phonological representations may be relatively good due to their age-appropriate reading skills, and children with SLI + RI, whose phonological representations are relatively weak because they cannot make use of the feedback from the orthographic system as much as typically reading children can. We further compared the results of the children with SLI with those of a group of children with RI only. Children with RI only generally have typical vocabulary development but less-developed phonological skills (see Vellutino, Fletcher, Snowling, & Scanlon, 2004). The current results give insight into the relative roles of vocabulary and phonological skills in the ability to use phonotactic frequency in NWR.

Research Rationale, Questions, and Predictions

The roles of ND and lexical frequency in word recognition and of PP in NWR were investigated in two studies to evaluate the lexical and sublexical processing profiles of children with SLI and/or RI. We compared their performances with the performances of two groups of TD children, one of which was matched on chronological age (CA TD) of the children with SLI and/or RI and one of which was matched on receptive vocabulary of the children with SLI (VOCAB TD). The latter group was approximately 2 years younger than the other groups.

In Study 1 we addressed our first research question, which was whether ND and lexical frequency influence word recognition in Dutch TD children and whether such an influence increases with age due to lexical growth. We expected that the older TD children would be less accurate in recognizing high-ND words compared with low-ND words.

We furthermore hypothesized that there would not be an effect of ND for the younger children because of their less developed lexicons. We further expected an interaction between lexical frequency and ND for the older TD children. The second research aim of Study 1 was to investigate the effect of ND and lexical frequency in children with SLI and/or RI. We hypothesized that the RI-only children would show the same negative effect of high ND on word recognition and the same lexical frequency effect as the CA TD children. For the SLI-only and the SLI + RI children, we expected that their performance on the word-recognition task would be comparable to that of the younger VOCAB TD group in terms of effect of ND and lexical frequency.

In Study 2 we administered an NWR task to address our third and fourth research questions. We investigated whether there was a developmental difference in the TD children with respect to the influence of PP on repeating nonwords and whether this was related to vocabulary. On the NWR task we expected the younger children to show a larger influence of low PP due to their smaller vocabularies. We further questioned whether the children with SLI were affected by PP to the same extent as their CA TD peers when they had to repeat nonwords. Based on previous research, we expected that children with a relatively small vocabulary size would have generally more problems repeating nonwords, especially with items low in PP, and that they would perform similarly to the younger VOCAB TD children. If the use of phonotactic frequency information is dependent on vocabulary, we expected the children with RI only to have a similar PP effect as the CA TD children, considering they have similarly developing lexicons. Researchers have also suggested that NWR is facilitated by robust phonological representations (Metsala, 1999; Rispens & Baker, 2012). As RI is generally associated with poor phonological skills (Elbro, Borstrøm, & Petersen, 1998; Vellutino et al., 2004), it can also be hypothesized that children with RI only would be affected more by low PP than the CA TD children and the children with SLI only.

Finally, the two experimental tasks were designed using the neighborhood counts, the lexical frequency counts, and the phonotactic frequency counts of the Dutch language. The results thus demonstrate whether the effects of these psycholinguistic variables that have been found in English are also observed in Dutch-speaking children. Dutch and English are both Germanic languages with almost the same syllable pattern, but there are differences in the phonotactic restrictions between the languages (Zamuner, 2006). The results of the two tasks establish whether the effects of ND, lexical frequency, and PP are not restricted to English but rather hold cross-linguistically.

Study 1: The Influence of ND and Lexical Frequency on Word Recognition

Method

Participants

Five groups of participants took part in this study. Informed consent was obtained from the parents of all participants. The participants have also been described by Rispens and Baker (2012), but that research concentrated on another question. There were three groups of children with language disorders: children with SLI only, children with SLI + RI, and children with RI only. Two groups of TD children participated: One group (CA TD; mean age 7;8) was matched on the basis of chronological age of the children with SLI and/or RI. The second group (VOCAB TD) was matched on the receptive vocabulary scores of the children with SLI only and SLI + RI using the vocabulary score generated by the Peabody Picture Vocabulary Test (PPVT; Dutch version: Schlichting, 2005). Table 1 gives an overview of the characteristics of all participants, who are discussed below.

TD children. The CA TD children (n = 38, mean age = 7;8) were selected from four primary schools located in the north, middle, and southwest regions of the Netherlands. All children were in second grade and were progressing normally at school. Their teachers confirmed that they had normal hearing, normal or corrected-to-normal vision, and no cognitive or emotional disturbances. As determined on the basis of teacher information, all children were from a family where at least one parent was a native speaker of Dutch and where Dutch was the language spoken at home from birth onward. The children were assessed with the Dutch version of the Raven Standard Progressive Matrices (Raven, 2006) for an estimation of their nonverbal intelligence quotient (IQ) performance at the time of testing. All children had a percentile score of at least 15, and the mean scores were within normal limits (see Table 1). The children were also administered two reading tasks (see next section for details of those tasks) and were assessed with the Dutch PPVT (Schlichting, 2005); see Table 1 for the results.

The VOCAB TD children (n = 16, mean age = 5;8) fulfilled the same criteria but were in the second year of kindergarten. The Raven task could not be administered to this group of children because there are no norm scores available for this age group in the Netherlands. The teachers of the children reported normal progress in kindergarten based on evaluation measures used in school to track children's progress. The children were administered the Dutch PPVT (Schlichting, 2005) for an indication of their receptive vocabulary; see Table 1 for the results.

Children with SLI and/or RI. Three subgroups of children with SLI and/or RI were formed. Thirty-three participants with a diagnosis of SLI participated (mean age = 8;0). The children with SLI were selected from three special-needs schools for children with developmental language disorders. The diagnosis of SLI had already been established by a multidisciplinary team at the schools because this is a requirement to be accepted into the schools. The team used either a cutoff of at least 1.5 *SD* below the mean in at least two language domains measured on a selection of Dutch standardized language tests or more than 2 *SD* below the mean on one Dutch standardized general language test. Children who had a history of speech output problems, such as dyspraxia, or showed evidence of such problems at the moment of testing, were excluded. All children had normal hearing and normal or

Table 1. The mean and standard deviations of age (years;months), reading scores, receptive vocabulary scores, Raven scores, and past tense production score (percentage accurate production of past tense) for five groups of children.

Variable	SLI + RI		SLI only		RI only		CA TD		VOCAB TD	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age	8;0	0;2	8;1	0;2	8;3	0;7	7;8	0;5	5;8	0;3
RWT	2.8[d]	1.9	10.2	1.9	4.9[d]	2.7	12.1	2.6	—	
PWT	4.3[d]	2.2	10.8	2.2	4.5[d]	1.5	11.9	2.9	—	
PPVT stand	86.0[a,b,c]	12.1	89.4[a,b,c]	10.2	100.5	8.25	106.2	9.9	103.4	18.2
PPVT raw	89.7[a,c]	10.2	92.4[a,c]	8.9	103.7	10.5	104.7	9.6	81.8[a,c]	16.7
Raven	41.3[a]	17.4	62.7	24.9	56.9	29.0	65.5	23.8	—	
Past tense	17.1[a]	27.8	23.1[a]	29.1	—		73.4	29.2	—	

Note. SLI = specific language impairment; RI = reading impairment; CA TD = typically developing children age matched on the chronological age of the children with SLI and/or RI; VOCAB TD = typically developing children matched on the receptive vocabulary of the children with SLI; RWT = real-word test, standard scores; PWT = pseudoword reading test, standard scores; PPVT stand = Peabody Picture Vocabulary Test (Dutch version: Schlichting, 2005), standard scores; PPVT raw = Peabody Picture Vocabulary Test (Dutch version), raw scores; Raven = Raven Standard Progressive Matrices (Raven, 2006). Dashes indicate data not available.

[a]Significantly lower than the CA TD group. [b]Significantly lower than the VOCAB TD group. [c]Significantly lower than the RI-only group. [d]Significantly lower than the SLI-only group. Significance $p < .05$.

corrected-to-normal vision. On the basis of teacher information, only children who were raised in a family with at least one parent who was a native speaker of Dutch and where Dutch had been spoken from birth onward were included. The special-needs schools were located in two regions (north and middle) of the Netherlands. The children with SLI who had been selected for our study were assessed with the Dutch version of the Raven Standard Progressive Matrices (Raven, 2006) for an estimation of their nonverbal IQ performance at the time of testing. All children with SLI had a percentile score of at least 15, and the mean score of the group was within normal limits (group of children with SLI: $M = 47.88$, $SD = 21.96$; see Table 1).

The children with SLI were screened for RI using two tasks: a real-word task (RWT; Brus & Voeten, 1973) and a pseudoword task (PWT; van den Bos, Spelberg, Scheepstra, & de Vries, 1994). In the RWT the child is required to read out a list of existing words as quickly and accurately as possible within 1 min. The PWT follows the same principle but uses nonwords and takes 2 min. Timed reading tasks were chosen because in the Netherlands, speed is judged to be a better indicator of reading development than accuracy alone due to the relative transparency of the orthography (de Jong & van der Leij, 2003). Both tasks are commonly used as part of the diagnosis procedure of developmental dyslexia. Reading impairment was established when children scored more than 1 *SD* below the mean on both the RWT and the PWT (normed scores: *M* standard score = 10, *SD* = 3). From the original group of 33 children with SLI, two groups were formed: 23 children were SLI + RI and the remaining 10 were SLI only (see Table 1). A one-way analysis of variance (ANOVA) was carried out to examine the differences between the children with SLI + RI and children with SLI only on the two reading tasks. Post hoc tests (Games-Howell due to unequal group sizes) were carried out as a follow-up. For the RWT there was a main significant effect of group, $F(3, 84) = 85.58$, $p < .001$, and the children with SLI + RI scored significantly lower than the children with SLI only and the TD children (both $p < .001$). The children with SLI only did not differ from the CA TD children ($p = .17$). For the PWT there was also a main effect of group, $F(3, 84) = 62.88$, $p < .001$. The children with SLI + RI had significantly lower scores than the children with SLI only and the CA TD children (both comparisons $p < .001$). The children with SLI only and the CA TD children did not differ on the PWT scores ($p = .57$).

It was important to check that the two groups of SLI children differed only with respect to their literacy scores but not on their general language ability, because this might otherwise confound results. Two language measures were compared between the two subgroups. One was the receptive vocabulary measure, the Dutch PPVT (Schlichting, 2005). An ANOVA was carried out to examine group comparisons. The children with SLI scored significantly lower than the CA TD children (see Table 1), but importantly there was no statistical difference between the SLI children with RI and without RI (raw score, $p = .45$; standard scores, $p = .67$). A past-tense production task had also been administered to the same children for a different study (Rispens & de Bree, 2014) and was considered a suitable background measure because problems with the production of the past tense are characteristic of SLI in Dutch and in many other languages (Bishop, 1997; de Jong, 1999; Leonard, 1998). It is important to note that there was no statistical difference between the two groups of children with SLI ($p = .22$), but both groups scored significantly lower than the CA TD children ($p < .001$). Table 1 presents the results of the groups of children with SLI and the CA TD children; no data are available for the VOCAB TD children and the children with RI only.

The children with RI but without any impairment in their spoken language ($n = 14$, mean age = 8;3) all attended regular schools and were contacted via special-needs teachers. Their diagnosis was based on scores at least 1 *SD* below the mean on standardized technical reading and spelling tests. The diagnosis was confirmed by the scores on the RWT and PWT administered in this study (see Table 1). The one-way ANOVA with follow-up tests (Games-Howell due to unequal sample sizes) described above showed that the group of children with RI only scored significantly lower on the two reading tasks compared with the CA TD children and the children with SLI only (RWT: both comparisons $p < .001$; PWT: both comparisons $p < .001$). The children with RI only did not significantly differently than the children with SLI + RI (RWT, $p = .23$; PWT, $p = .99$). Teachers and parents informed us that the children with RI had never been diagnosed as having problems in their oral language and had never been referred to a speech and language therapist for such problems. Their receptive vocabulary scores (i.e., PPVT) were similar to those of the CA TD children (see Table 1). The children with RI were assessed with the Dutch version of the Raven Standard Progressive Matrices (Raven, 2006) for an estimation of their nonverbal IQ performance at the time of testing. All children had a percentile score of at least 19, and the mean scores were within normal limits (see Table 1).

The CA TD children were matched on chronological age to the children with SLI and/or RI and the children with RI only. An ANOVA showed that there were no differences in age between the children, but the age difference between the RI-only children and the CA TD children approached significance, with the RI-only children being older than the CA TD children (Games-Howell post hoc tests due to unequal group sizes: CA TD vs. SLI only, $p = .42$; CA TD vs. SLI + RI, $p = .18$; CA TD vs. RI only, $p = .08$; SLI only vs. SLI + RI, $p = 1.0$; SLI only vs. RI only, $p = .59$; SLI + RI vs. RI only, $p = .72$).

Word Recognition: Lexical Decision Task

A lexical decision task was constructed to measure word recognition. This task has also been reported by Rispens and Baker (2012) but without any discussion of the variable ND, lexical frequency, or the effect of ND and lexical frequency on the task performance.

Forty-eight target words were selected, and 48 minimally deviating nonwords were constructed in order to test accuracy of word recognition of the lexical targets in a picture-to-word matching design. The minimal word–nonword pairs were presented auditorily. The nonword item of the pair differed in one phoneme (always a consonant) from the word. An example of a word pair is *tijger–dijger* (*tiger–diger*). In half of the items, the consonant manipulated to generate a minimally deviant nonword was in the onset position of the initial stressed syllable (e.g., *tiger–**d**iger*). In the other half of the items, the target phoneme was in a medial position. This was the onset position of the second unstressed syllable (e.g., *motor–mo**p**or*).

All real-word items consisted of two syllables and followed the dominant stress pattern of Dutch (i.e., first syllable stressed). All real words were nouns, were highly imageable, and had an early age of acquisition (all items below the age of 6 years; Schaerlaekens, Kohnstamm, & Lejaegere, 1999).

The main variables of interest were high versus low ND and high versus low frequency. ND was defined here as the number of words that differed from the target word by one phoneme substitution, addition, or deletion while preserving relative position of the segments (Charles-Luce & Luce, 1995). ND was determined using the Dutch version of WordGen (Duyck, Desmet, Verbeke, & Brysbaert, 2004). The median of the ND values of the selected words was calculated (*Mdn* = 4), and the division between high and low ND was based on a median split ±1 neighbor. Items with an ND value +1 of the median were considered high in ND (ND between 5 and 14), and items with an ND value −1 of the median were considered low in ND (ND between 0 and 3). ND and lexical frequency were orthogonally combined into four sets: 12 words with high lexical frequency and high ND, 12 word with high lexical frequency and low ND, 12 words with low lexical frequency and high ND, and 12 words with low lexical frequency and low ND. The lexical frequency count was based on the CELEX database (Baayen, Piepenbrock, & Gulikers, 1995). Half of the nouns used were low in lexical frequency (log frequency 0–1.18, $M = 0.51$, $SD = 0.33$). The other half consisted of high-frequency words (log frequency 1.58–2.68, $M = 1.99$, $SD = 0.39$). A t test showed that the frequency of the items with high lexical frequency was significantly higher than that of the low-frequency items, $t(46) = 14.0$, $p < .001$. The difference between items high and low in ND was statistically significant, $t(46) = 13.14$, $p < .001$ (see Appendix A for all items and the ND and lexical frequency values). The size of the neighbors may be lower than that in previous studies (see, e.g., Mainela-Arnold et al., 2008), but this may be caused by the inclusion of disyllabic words instead of single-syllable words, considering short words tend to have higher ND than longer words (Storkel, 2004).

In addition, the PP of the nonword counterpart of the word–nonword pair was calculated. The PP of the incorrect items was controlled for to make sure that the PP of the nonwords could not interfere with the ability to process the nonword. The PPs of the nonwords were calculated using the Dutch phonotactic frequency database (Adriaans, 2006) derived from the corpus of spoken Dutch (Oostdijk, 2000). The total PP of all sequences of phonemes per nonword was calculated, starting with the probability of a phoneme in the word-initial position, followed by the probability of the combination of all sequential biphones, and ending with the probability of the phoneme in the word-final position. The total PP was then divided by the number of biphones to calculate the mean: The closer to zero, the higher the PP. Appendix A lists all mean PPs per nonword (range = −0.97 to −1.55). To investigate whether the mean PP was different between the high- and low-ND nouns, the nonword items were grouped on the basis of the ND category of the noun counterparts (i.e., all nonword items belonging to a pair in which the noun was high ND were grouped together; likewise with low-ND nouns). We checked whether the PPs of the nonword counterparts of the high- and low-ND items differed in PP. A t test showed that this was not the case, $t(46) = 1.16$, $p = .25$. Thus, the nouns of the recognition test were varied on ND and lexical frequency, and the nonword counterparts were controlled for PP. Six filler items were presented twice in the correct form (see Appendix A for all items).

All items were prerecorded by a female native speaker of Dutch and were digitized. Items were auditorily presented from a laptop computer using the software E-prime

(Psychology Software Tools). This software program also recorded the responses of the children.

Procedure and Data Analysis

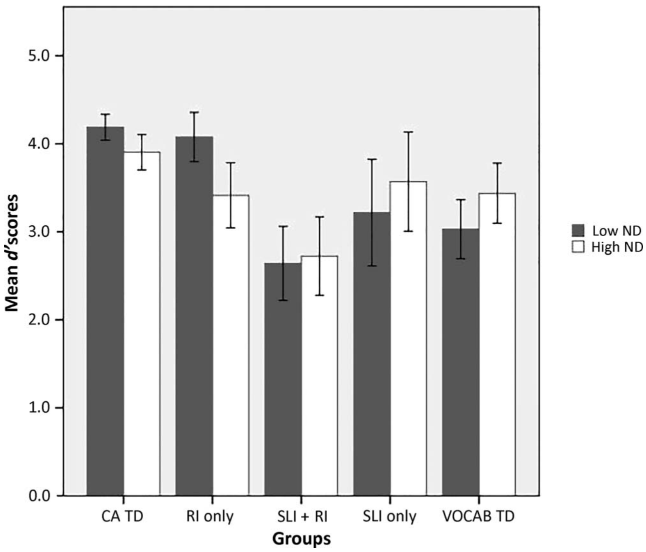
All children were seen individually in a separate room at their schools or, on a few occasions, at their homes. The task was administered by trained speech and language therapists and speech and language intervention students. This task was administered together with other tasks on two separate test sessions that lasted around 45 min each. Children were shown a picture on a laptop computer and were then played the question "*is dit een . . .*" ("*is this a . . .*") followed by either a word or a nonword. All pictures were presented twice in a pseudorandomized order throughout the task, so that one picture elicited once a *yes* answer (correct word that matched the picture; e.g., the word *tiger* matching with a picture of a tiger) and once a *no* answer (nonword differing in one phoneme from the picture; e.g., the nonword *diger* presented with a picture of a tiger). All minimal pair members were presented once as the derived nonword form (e.g., *diger*) and once as the actual lexical word (e.g., *tiger*). Children had to press a button on the laptop computer to indicate their response to the question, and the responses were recorded by the software program. All responses were classified as *hits* (pressing *yes* in case of a matching lexical target), *misses* (pressing *no* in case of a matching lexical target), *false alarms* (pressing *yes* in case of a nonword), or *correct rejections* (pressing *no* in case of a nonword). We used signal detection analysis to calculate the mean d' score (see Green & Swets, 1966). This means that for each word–nonword pair, we calculated whether the existing word elicited the *yes* answer and the nonword elicited the *no* answer. The formula for calculating the d' is to subtract the z-transformed scores of the hits and false alarms: $d' = Z(\text{hit rate}) - Z(\text{false alarm rate})$. In the present study, a perfect d' sensitivity score was 4.65.

An ND (two levels: high and low) × lexical frequency (two levels: high and low) × group (five levels: CA TD, VOCAB TD, SLI + RI, SLI only, RI only) repeated measures ANOVA was carried out. In case of interaction, post hoc tests were carried out to follow up the interaction. For the group comparisons, Games-Howell tests were used due to unequal group sizes. The p value was set to < .05.

Results

For the overall performance on the task, there was a significant main effect of group, $F(4, 96) = 21.58$, $p < .001$, $\eta_p^2 = .47$ (see Figure 1). Post hoc tests showed that the CA TD children significantly outperformed all other groups ($p < .01$, $d > 1.50$) except for the RI-only children, where there was no significant difference ($p = .14$, $d = 0.74$). The SLI + RI children performed significantly lower than all other groups (SLI + RI vs. SLI only, $p = .048$, $d = -0.92$; SLI + RI vs. CA TD, $p < .001$, $d = -1.92$; SLI + RI vs. RI only, $p < .001$, $d = -1.50$) except for the VOCAB TD group ($p = .15$, $d = -0.74$). For the SLI-only group, no significant differences were found between the RI-only group ($p = .34$, $d = -0.83$) and the VOCAB TD group ($p = .91$, $d = 0.35$).

Figure 1. Mean d' scores of the items on the lexical decision task with low and high neighborhood density (ND). SLI = specific language impairment; RI = reading impairment; CA TD = typically developing children age matched on the chronological age of the children with SLI and/or RI; VOCAB TD = typically developing children matched on the receptive vocabulary of the children with SLI. Error bars represent 95% confidence intervals.



The RI-only group significantly outperformed the VOCAB TD group ($p = .03$, $d = 1.16$).

No significant main effects of lexical frequency, $F(1, 96) = 1.06$, $p = .31$, $\eta_p^2 = .011$, and ND, $F(1, 96) = 0.16$, $p = .69$, $\eta_p^2 = .002$, were found, nor was there a significant interaction between lexical frequency and ND, $F(1, 96) = 0.02$, $p = .78$, $\eta_p^2 = .05$. The interaction between lexical frequency and group was not significant, $F(4, 96) = 1.10$, $p = .36$, $\eta_p^2 = .04$, in contrast with the interaction between ND and group, $F(4, 96) = 4.00$, $p = .005$, $\eta_p^2 = .14$. Follow-up investigations showed this interaction reflected that both the CA TD group and the RI-only group demonstrated significant main effects of ND (CA TD, $p = .02$, $\eta_p^2 = .14$; RI only, $p = .016$, $\eta_p^2 = .41$), with both groups showing significantly higher d' scores for the low-ND items (see Figure 1). The VOCAB TD group ($p = .14$, $\eta_p^2 = .14$), SLI-only group ($p = .48$, $\eta_p^2 = .07$), and SLI + RI group ($p = .54$, $\eta_p^2 = .012$) did not show a significant effect of ND; in fact, the ND effect seemed reversed. (Note the large effect size for the VOCAB TD group and the medium effect size for the SLI-only group, even though it was nonsignificant.)

Discussion

Our first research question was whether TD children showed an effect of ND and lexical frequency on word recognition and whether developmental differences could be observed. As predicted, the older TD children showed a significant negative effect of high ND, whereas the younger TD children did not show a significant effect of ND. In fact, even though it was not significant, the effect of ND seemed to be reversed in the latter group, as the scores on the high-ND items were higher than the scores on the low-ND items. Thus, our study underlines Garlock et al.'s (2001) findings that the effect of ND is dependent on lexical development. It is important to note that our results show that such findings are not bound to the English language but rather also hold for Dutch-speaking children. The results further show that lexical frequency did not significantly influence task performance and that it did not interact with any of the other variables. These results are similar to those of Garlock et al., who argued that age of acquisition may be a more sensitive index of lexical familiarity than lexical frequency. Because we kept age of acquisition constant in our study, this may have caused the absence of the frequency effect.

Our second research question asked whether children with SLI and/or RI were influenced by ND and lexical frequency to the same extent as the TD children. First, none of the groups showed an effect of lexical frequency or an interaction between ND and lexical frequency. This was similar to the TD groups. We expected that the children with RI only would show effects similar to those of ND as the CA TD children. The findings supported our hypothesis. These results are consistent with those of Griffiths and Snowling (2001) and those of Thomson, Richardson, and Goswami (2005), who demonstrated that in a working memory task (serial recall) children with RI were affected by ND to the same extent as their age-matched TD peers. The results are inconsistent with those of Metsala (1997). This is possibly caused by the differences in receptive vocabulary between the children with RIs and the control children in her study (with the control children having higher vocabulary scores). This was not the case in our study and in that of Griffiths and Snowling. Our findings thus confirm the suggestion of Griffiths and Snowling (2001) that RI children who have normal vocabulary skills and who do not show concomitant language difficulties do not have problems with word recognition and show the same effect of ND.

For the groups of children with SLI (and/or RI), we hypothesized that they would behave similarly to the younger TD children who were matched on their vocabulary scores. This prediction was supported as both the SLI groups and the VOCAB TD group did not show significant effects of ND and of lexical frequency. Our results thus support the hypothesis of Griffiths and Snowling (2001) that children with reading disabilities who also show oral language weaknesses are affected in word recognition.

The present results further suggest that the organization of the mental lexicon in children with SLI is delayed, considering their results—in terms of both accuracy and effect of ND—were comparable to those of TD children 2 years younger who have similar receptive vocabulary skills. The current results (combining those of the clinical groups and the TD groups) thus suggest that the lexicon needs to be of a substantial size to generate a competition effect and that the effect of ND on word recognition is related to vocabulary. As noted in the introduction, Mainela-Arnold et al. (2008) did not find a difference in ND effect between children with SLI and TD children. This is different from our results. The difference in age between groups may have caused the difference between the two studies. The children with SLI in the Mainela-Arnold et al. study were on average 2 years older, and it can be estimated that their vocabularies were larger compared with those of the children in our study. The absence of the ND effect in our sample in comparison with the presence of the ND effect in their study may thus reflect a difference in lexical development. An alternative explanation may be the differences between the tasks used (word recognition in a lexical decision task vs. a gating task). The current task involved a picture so that the lexical–semantic level of a lexical item was activated. For children with weak lexical–semantic processing (generally the case in SLI), such activation places extra demands on the processing involved in the task, which may have interfered with their performance.

Study 2: Influence of PP on NWR and Its Association With Vocabulary

Method

Participants

All the TD children who participated in the previous study also participated in this study, but additional children were included using the same selection criteria. The CA TD group from Study 1 was enlarged to 80 children; all were attending the second grade of elementary school and were 7, 8, or 9 years old (n = 80, mean age = 7;11, 48 girls). The VOCAB TD group was also enlarged (n = 48, mean age = 5;10, 23 girls); all children attended kindergarten.

The same children with SLI and/or RI who participated in Study 1 also participated in this study. Informed consent was obtained from the parents of all participants.

All children were assessed with the Dutch version of the PPVT (Schlichting, 2005); see Table 2 for the results. The TD kindergarten children did not significantly differ from the children with SLI only and the children with SLI + RI with respect to receptive vocabulary and are referred to as VOCAB TD. The CA TD children significantly outperformed the children with SLI only and the children with SLI + RI ($p < .016$). A one-way ANOVA showed that there was no significant difference between the ages of the age-matched groups, even though the p value is rather low, $F(3, 126) = 2.37$, $p = .074$. Post hoc tests (Games-Howell) indicated that none of the group comparisons with respect to age were significant ($p > .23$).

NWR Task

The NWR task consisted of 40 items that were divided equally into two-, three-, four-, and five-syllable items. As in the previous word-recognition study, the PP of the items was calculated on the basis of the Dutch phonotactic frequency database (Adriaans, 2006) derived from the corpus of spoken Dutch (Oostdijk, 2000). The total PP of all sequences of phonemes per nonword was calculated, starting with the probability of a phoneme in the word-initial position, followed by the probability of the combination of all sequential biphones, and ending with the probability of the

Table 2. Mean performance and standard deviations on the Peabody Picture Vocabulary Test (PPVT; Dutch version: Schlichting, 2005; raw score) and the phoneme percentage correct on the nonword repetition (NWR) task.

Task	SLI + RI		SLI only		RI only		CA TD		VOCAB TD	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
PPVT raw	89.5[a,c]	10.4	92.4[a,c]	8.9	103.7	9.0	104.5	9.6	88.6[a,c]	13.2
NWR total	72.9[a,b,c,d]	5.9	82.8	5.9	85.3	5.9	88.6	4.9	85.2[a]	6.8

Note. SLI = specific language impairment; RI = reading impairment; CA TD = typically developing children age matched on the chronological age of the children with SLI and/or RI; VOCAB TD = typically developing children matched on the receptive vocabulary of the children with SLI. [a]Significantly lower than the CA TD group. [b]Significantly lower than the VOCAB TD group. [c]Significantly lower than the RI-only group. [d]Significantly lower than the SLI-only group. Significance $p < .05$.

phoneme in the word-final position. The PPs are all negative numbers; the closer the PP is to zero, the more frequent the biphone. The mean PP (the total PP divided by the number of biphones) of the nonword was categorized as either low ($M = -1.95$, $SD = 0.16$) or high ($M = -1.31$, $SD = 0.09$; see Appendix B for a list of all items and their PPs). Items of high and low PP were distributed equally over the four word lengths so that in each word length condition five items were high in PP and five were low in PP. A *t* test confirmed that the PPs of the low items and the PPs of the high items were significantly different from each other, $t(40) = 15.02$, $p < .001$. Because there were four length conditions, we also carried out *t* tests for each word length to check whether there was a difference in the mean PPs of the low and high conditions for each word length condition. All four tests showed that the mean PP of the items classified as low PP was significantly lower than the mean PP of items classified as high PP—two syllables: low PP, $M = -1.80$, $SD = 0.06$, vs. high PP, $M = -1.19$, $SD = 0.03$, $t(8) = 21.32$, $p < .001$; three syllables: low PP, $M = -1.97$, $SD = 0.25$, vs. high PP, $M = -1.28$, $SD = 0.07$, $t(8) = 5.78$, $p < .001$; four syllables: low PP, $M = -2.06$, $SD = 0.06$, vs. high PP, $M = -1.41$, $SD = 0.08$, $t(8) = 15.13$, $p < .001$; five syllables: low PP, $M = -1.97$, $SD = 0.11$, vs. high PP, $M = -1.36$, $SD = 0.04$, $t(8) = 11.68$, $p < .001$.

The items did not contain consonant clusters. All syllables were consonant–vowel except for the word-final syllable, which was consonant–vowel–consonant. All items followed the regular Dutch stress pattern. Note that this task has also been discussed in Rispens and Baker (2012) but without any reference to the variable PP.

Procedure and Data Analysis

All children were seen individually at their schools or homes in a quiet location. The PPVT was administered first, followed by the NWR task. The NWR started with an explanation of the task followed by three practice items of two, three, and four syllables. The test items were prerecorded by a female native speaker of Dutch and were played back using a laptop computer with external loudspeakers. The items were pseudorandomly divided into two blocks of 20 items each and were presented to the children with a pause between the blocks. The order of the two blocks was pseudorandomized. All responses of the children were recorded and transcribed offline. Omissions and substitutions of target phonemes were counted as errors. For each item the percentage of phonemes correctly repeated per item (percentage phonemes correct; PPC) was calculated because this scoring method is more sensitive than scoring the number of items repeated accurately (Graf Estes, Evans, & Else-Quest, 2007). All items were transcribed and scored by at least two trained researchers who were blind to the participant groups. All scores were compared. In case of differences in scoring, the raters discussed the transcription and scoring together and made a decision.

The PPC of all participants was analyzed on group differences, on differences between high and low PP, and on a possible interaction between group and PP. A repeated measures ANOVA was used with group (five levels: CA TD, VOCAB TD, SLI only, SLI + RI, and RI only) as the between-subjects variable and PP (two levels: low and high) as the within-subject variable. Post hoc tests were carried out to examine main effects (using Games-Howell due to unequal sample sizes). The *p* value was set to < .05.

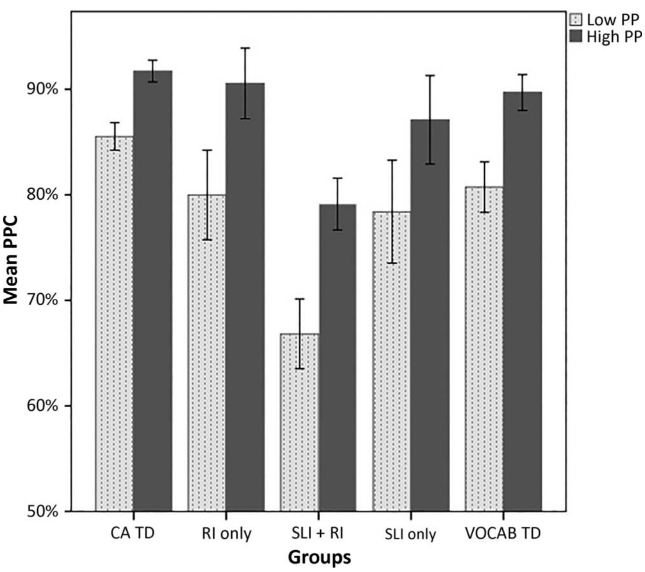
Results

Table 2 displays the results on the NWR and PPVT. The PPC on the NWR task was divided into the overall score on the low-PP and high-PP items (see Figure 2).

Influence of PP on NWR

The repeated measures ANOVA showed that there was a significant main effect of PP, $F(1, 171) = 320.41$, $p < .001$, $\eta_p^2 = .65$; a significant main effect of group, $F(4, 171) = 32.96$, $p < .001$, $\eta_p^2 = .44$; and a significant interaction between group and PP, $F(4, 171) = 7.36$, $p < .001$, $\eta_p^2 = .15$. The main effect of group reflected that the children with SLI + RI performed overall significantly below all other groups ($p < .002$, $d < -1.75$). The children with SLI only performed comparably to all other groups (SLI only vs. RI only, $p = .83$, $d = -0.44$; SLI only vs. VOCAB TD, $p = .75$, $d = -0.39$; SLI only vs. CA TD, $p = .066$, $d = -1.10$, but note the *p* value and the large effect size) except for the children with SLI + RI, whom they significantly outperformed ($p = .002$, $d = 1.75$). The children with RI did not perform significantly differently from both TD groups (RI only vs. CA TD, $p = .35$, $d = -0.59$; RI only vs. VOCAB TD, $p = 1.0$, $d = 0.001$) or from the children with SLI only ($p = .83$, $d = 0.35$). They significantly

Figure 2. Mean phoneme percentage correct (PPC) of the items of the nonword repetition task with low and high phonotactic probability (PP). SLI = specific language impairment; RI = reading impairment; CA TD = typically developing children age matched on the chronological age of the children with SLI and/or RI; VOCAB TD = typically developing children matched on the receptive vocabulary of the children with SLI. Error bars represent 95% confidence intervals.



outperformed the children with SLI + RI ($p < .001$, $d = 2.10$). The CA TD children outperformed both the group of children with SLI + RI ($p < .001$, $d = 2.95$) and the younger VOCAB TD children ($p = .035$, $d = 0.53$).

The main effect of PP reflected that overall the items with high PP were repeated significantly more accurately than items with low PP (see Figure 2). Follow-up tests to examine the interaction between group and PP showed that the children with SLI + RI were significantly more affected by low PP than both control groups (CA TD, $p < .001$, $\eta_p^2 = .19$; VOCAB TD, $p = .019$, $\eta_p^2 = .077$). In contrast, the influence of PP for the SLI-only group was not significantly different from that of both control groups (CA TD, $p = .14$, $\eta_p^2 = .025$; VOCAB TD, $p = .88$, $\eta_p^2 = 0$). The group of children with RI only was significantly more affected by low PP compared with the CA TD group ($p = .003$, $\eta_p^2 = .098$) but demonstrated a pattern similar to that of the VOCAB TD group ($p = .30$, $\eta_p^2 = .018$). The VOCAB TD group was significantly more affected by low PP than the CA TD group ($p < .001$, $\eta_p^2 = .071$). The effect of PP was not significantly different between the children with SLI only, children with SLI + RI, and children with RI only (SLI only vs. SLI + RI, $p = .15$, $\eta_p^2 = .065$; SLI only vs. RI only, $p = .45$, $\eta_p^2 = .026$; SLI + RI vs. RI only, $p = .43$, $\eta_p^2 = .018$).

Relation Between NWR and Receptive Vocabulary in TD Children

A correlation analysis was carried out to investigate the relationship between receptive vocabulary and NWR performance. The PPVT score was entered together with the low-PP and high-PP items of the NWR for each group separately. For the kindergarten group the low-PP (Pearson's $r = .42$, $p = .002$) and high-PP (Pearson's $r = .31$, $p = .03$) items significantly correlated with receptive vocabulary (correlation PPVT and mean NWR, $p = .005$). For the second-grade group no significant correlations were found (Pearson's $r = .20$, $p = .07$; high-PP items, Pearson's $r = .14$, $p = .21$; mean NWR, $p = .08$). These results suggest that the association between vocabulary size and NWR decreases over time. During kindergarten vocabulary correlates with NWR, whereas in the older group items with a high PP do not correlate with vocabulary size and items with a low PP are not, or are only marginally, significantly associated with vocabulary size.

Relation Between NWR and Receptive Vocabulary in Children With SLI and RI

The association between vocabulary and NWR performance within the individual groups of SLI-only, SLI + RI, and RI-only children was also analyzed using a nonparametric correlation analysis (Spearman's rho) due to the skewed distribution of the data. No significant correlations were found for each group separately ($p > .5$).

Discussion

Our third research question addressed whether the influence of PP on NWR during typical development and the association with vocabulary changed over time. Our results showed that the influence of PP and the association between vocabulary and NWR performance was different for the two TD groups. Even though in both groups of TD children there was a significant effect of PP, in the sense that low-PP items were repeated significantly less accurately than high-PP items, it was found that the effect of PP was significantly stronger for the younger TD children. These results not only confirm the findings of Edwards et al. (2004) but also strengthen the evidence on the influence of PP on NWR, because these results give cross-linguistic support. The correlation analysis furthermore supported the previous findings of Munson et al. (2005) that this decrease of the PP effect over time is associated with vocabulary growth. NWR performance was significantly correlated with vocabulary size in the kindergarten TD children, whereas this was not the case for the older TD children.

Finally, our fourth question investigated whether children with SLI and/or RI were affected to the same extent as TD children by PP on an NWR task. Previous studies rendered discrepancies with regard to that effect. Our results showed that the groups of children with SLI and/or RI demonstrated a significant effect of PP, with low PP resulting in less accurate repetitions. However, the three groups differed with respect to the extent of this effect relative to the TD groups. The children with RI only demonstrated a significantly larger (negative) effect of low PP than the CA TD group and performed similarly to the younger VOCAB TD group. This difference cannot be explained by a difference in receptive vocabulary because the RI-only and CA TD groups had comparable vocabulary scores. No significant correlations between vocabulary and NWR in the groups with language impairments were found. Rather, the results indicate that children with RI have difficulties in extracting phonological regularities from the lexicon when phonotactic patterns that are not frequently occurring are involved.

The effect of low PP was significantly larger for the SLI + RI group compared with the CA TD group, the VOCAB TD group, and the group of children with SLI only even though there were no significant differences found between their vocabulary scores and the vocabulary scores of the latter two groups. In contrast, the group of children with SLI only demonstrated a PP effect comparable with that of the CA TD group. This suggests that children with SLI only can flexibly manipulate phonological information in a task when they have to create a novel phonological representation. This ability is related to their reading ability because orthographic skills promote phonological awareness (see also Bishop et al., 2009; Rispens & Parigger, 2010).

The children with SLI + RI were affected in the same way by low PP as the RI-only children. This suggests that children with reading difficulties have problems with abstracting phonotactic information from phonological representations to facilitate the creation and maintenance of novel phonological representations, as demanded in NWR. However, it is not the case that children with RI only performed poorly overall on NWR. It is thus the effect of PP that is striking for the group of children with RI. This is different for the children with SLI + RI: This group showed an overall profound difficulty with NWR. The combination of a relatively small vocabulary and difficulty with abstracting phonotactic patterns from phonological representations may have resulted in this considerable difficulty with the task compared with the children with RI only, who have vocabulary scores appropriate for their age.

General Discussion

Two experimental tasks were administered to gain insight into processing at the lexical and sublexical levels in children with and without language and/or reading disorders. To this end, the effect of ND and lexical frequency on word recognition and the effect of PP on the ability to repeat nonwords were measured.

No effects of lexical frequency were found, but clear developmental effects were demonstrated for both ND and PP. Only the older TD children (7- to 8-year-olds) were significantly affected by high ND in word recognition, reflecting competition between the neighbors of the target word. This was not observed in the 5- to 6-year-olds. The older children have larger vocabularies (reflected by better receptive vocabulary scores), which implies that more lexical neighbors are activated when the target word is presented. Inhibition of such activation is at the expense of recognition accuracy. On the other hand, the older TD children had fewer problems with low-PP items when repeating nonwords compared with the younger children. Because of the larger vocabulary size, more lexical representations (including words with less frequent phonotactic sequences) are available that facilitate the recall of a nonword (see also Edwards et al., 2004; Metsala, 1999; Munson et al., 2005).

The groups with SLI and/or RI revealed some interesting differences. The RI-only group was not significantly different from the CA TD group in terms of the ND effect, and both groups performed similarly on the word-recognition task. We can therefore conclude that the organization of the mental lexicon of children with RI without concomitant language difficulties is similar to that of typically reading children, supporting previous findings of Griffiths and Snowling (2001). However, in the current study, despite a mean score on NWR comparable to the CA TD group, the children with RI only were significantly more affected by low PP than the CA TD children. It is important to note that this cannot be the result of differences between the vocabulary skills of both groups. Several studies have suggested that phonological encoding is particularly weak in children with RI (Griffiths & Snowling, 2001; Vellutino et al., 2004). Ramus and Szenkovits (2008) proposed that the phonological representations themselves are not deficient but rather that accessing phonological representations is impaired, especially during tasks taxing phonological short-term capacity, which is the case in NWR (see also Griffiths & Snowling, 2001). The current results of the combined tasks give evidence for this view on the phonological deficit in RI. The ND effect indicates that word recognition in RI only is unimpaired, and it indicates that the children with RI in our sample had relatively well-specified phonological representations; otherwise, the ND effect would not have been found. However, accessing and abstracting relatively infrequent phonotactic sequences from existing words in the mental lexicon in circumstances where the demands on phonological short-term memory are high was found to be problematic for these children. This was reflected in the lower scores on the low-PP items. These data thus indicate that even though NWR is associated with vocabulary, vocabulary itself does not predict the extent of the PP effect.

Compared with the children with RI only, the children with SLI only showed the opposite pattern. As a group, they showed no significant effect of ND, likely caused by their relatively low vocabulary scores. The effect of PP on NWR was comparable to that of the CA TD children. These findings thus suggest impaired lexical processing but normal sublexical processing. This latter finding fits the hypothesis of Bishop et al. (2009), which states that, as a result of adequate orthographic skills, children with SLI only have phonological abilities that are similar to those of TD children and that these abilities facilitate NWR.

The children with SLI + RI demonstrated the lowest scores of all groups. They did not demonstrate an effect of ND, reflecting an underdeveloped organization of the mental lexicon. This supports the idea of Griffiths and Snowling (2001) that only those children with RI who also show language difficulties are affected in word recognition. The negative effect of low PP was prominent in this group, indicating that in addition to affected lexical processing they have problems with accessing and using subsegments (especially those that involve infrequent phonotactic patterns) from the lexical representations in their mental lexicon (see also Coady et al., 2013). This group thus showed impairments at both the lexical and sublexical levels.

These findings demonstrate the importance of assessing reading skills when studying children with SLI, as sublexical processing was found to be associated with reading ability. They further indicate that vocabulary skills alone do not predict sublexical processing, because children with normal vocabulary scores (RI-only children) were more influenced by low PP than the CA TD children and children with relatively low vocabulary skills (SLI only) demonstrated a similar effect of PP compared to the CA TD children.

Finally, our findings did not show an effect of ND in children with SLI. In contrast, Mainela-Arnold et al. (2008) reported an effect of ND in 8- to 12-year-old children with SLI that did not differ from age-matched TD children. The combined results thus suggest that children with SLI are initially delayed in lexical processing but that an effect of ND will arise at a later developmental stage, presumably when their vocabularies are large enough. It also may be the case that our task was more difficult for the children with SLI compared with the gating task used in Mainela-Arnold et al.'s study, considering it involved activation of lexical–semantic information. Children with SLI often experience weaker lexical–semantic processing in comparison with their age-matched TD peers. A task that activates lexical–semantic information may thus cause more problems with lexical processing, which consequently can interfere with the effect of ND compared with a gating task that taps only the phonological level. Future studies are warranted to disentangle the potential effects of task demands, vocabulary skills, and lexical–semantic skill on word recognition in children with SLI.

The results from our studies could be strengthened in several ways. First, increasing the number of children in the groups with a language and/or reading disorder and matching them as closely as possible on chronological age with the CA TD children would improve statistical power. Second, investigating lexical and sublexical processing in children with SLI with and without RI would benefit from a longitudinal approach. This would reveal whether literacy acquisition aids the development of phonological representations and phonological processing or whether children with SLI who are able to learn to read and spell also show differences in processing phonological information before the onset of literacy acquisition. Furthermore, future research should include a range of tasks, as the methodology may affect to what extent variables impact lexical and sublexical processing.

Acknowledgments

This research was conducted as part of a Netherlands Organisation for Scientific Research Veni research grant, awarded to the first author. The authors would like to thank all children and their parents for participating in this study and the staff at their schools for their help. We furthermore thank Wendy Boelhouwer, Christa Kerkhof, and Martine Jong for their assistance in collecting the data.

References

Adriaans, F. (2006). PhonotacTools (Test version) [Computer program]. Utrecht, the Netherlands: Utrecht Institute of Linguistics OTS, Utrecht University.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX Lexical Database (Release 2, Dutch Version 3.1) [CD-ROM]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Bishop, D. V. M. (1997). *Uncommon understanding*. London, United Kingdom: Psychology Press.

Bishop, D. V. M., Hayiou-Thomas, M. E., McDonald, D., & Bird, A. (2009). Children who read words accurately despite language impairment: Who are they and how do they do it? *Child Development, 80,* 593–605.

Bishop, D. V., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin, 130,* 858–886.

Brus, B. T., & Voeten, M. J. M. (1973). *Een-Minuut Test. Vorm A en B*. Nijmegen, the Netherlands: Berkhout Testmateriaal.

Catts, H., Adlof, S. M., Hogan, T. P., & Ellis Weismer, S. (2005). Are specific language impairment and dyslexia distinct disorders? *Journal of Speech, Language, and Hearing Research, 48,* 1378–1396.

Charles-Luce, J., & Luce, P. A. (1995). An examination of similarity neighborhoods in young children's receptive vocabularies. *Journal of Child Language, 22,* 727–735.

Coady, J. A., Evans, J. L., & Kluender, K. R. (2010). The role of phonotactic frequency in nonword repetition by children with specific language impairments. *International Journal of Language & Communication Disorders, 45,* 494–509.

Coady, J. A., Mainela-Arnold, E., & Evans, J. L. (2013). Phonological and lexical effects in verbal recall by children with specific language impairments. *International Journal of Language & Communication Disorders, 48,* 144–159.

Criddle, M. J., & Durkin, K. (2001). Phonological representation of novel morphemes in children with SLI and normally developing children. *Applied Psycholinguistics, 22,* 363–382.

de Jong, J. (1999). *Specific language impairment in Dutch: Inflectional morphology and argument structure* (Doctoral dissertation). Retrieved from http://dissertations.ub.rug.nl/

de Jong, P. F., & van der Leij, A. (2003). Developmental changes in the manifestation of a phonological deficit in dyslexic children learning to read a regular orthography. *Journal of Educational Psychology, 95,* 22–40.

Duyck, W., Desmet, T., Verbeke, L., & Brysbaert, M. (2004). WordGen: A tool for word selection and non-word generation in Dutch, German, English, and French. *Behavior Research Methods, Instruments, & Computers, 36,* 488–499.

Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research, 47,* 421–436.

Edwards, J., & Lahey, M. (1998). Nonword repetitions of children with specific language impairment: Exploration of some explanations for their inaccuracies. *Applied Psycholinguistics, 19,* 279–309.

Elbro, C., Borstrøm, I., & Petersen, D. K. (1998). Predicting dyslexia from kindergarten: The importance of distinctness of phonological representations of lexical items. *Reading Research Quarterly, 33,* 36–60.

Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-acquisition, word frequency and neighborhood density effects on spoken word recognition: Implications for the development of phoneme awareness and early reading ability. *Journal of Memory and Language, 45,* 468–492.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics, 27,* 513–543.

Goh, W. D., Suárez, L., Yap, M. J., & Tan, S. H. (2009). Distributional analyses in auditory lexical decision: Neighborhood density and word frequency effects. *Psychonomic Bulletin & Review, 16,* 882–887.

Goswami, U. (2000). Phonological representations, reading development and dyslexia: Towards a cross-linguistic theoretical framework. *Dyslexia, 6,* 133–151.

Graf Estes, K., Evans, J. L., & Else-Quest, N. M. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 50,* 177–195.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: Wiley.

Griffiths, Y. M., & Snowling, M. J. (2001). Auditory word identification and phonological skills in dyslexic and average readers. *Applied Psycholinguistics, 22,* 419–439.

Leonard, L. B. (1998). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19,* 1–36.

Maillart, C., Schelstraete, M., & Hupet, M. (2004). Phonological representations in children with SLI: A study of French. *Journal of Speech, Language, and Hearing Research, 47,* 187–198.

Mainela-Arnold, E., Evans, J. L., & Coady, J. A. (2008). Lexical representations in children with SLI: Evidence from a frequency manipulated gating task. *Journal of Speech, Language, and Hearing Research, 51,* 381–393.

Mainela-Arnold, E., Evans, J. L., & Coady, J. (2010). Beyond capacity limitations II: Effects of lexical processes on word recall in verbal working memory tasks in children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 53,* 1656–1672.

McArthur, G. M., Hogben, J. H., Edwards, S. M., Heath, S. M., & Mengler, E. D. (2000). On the "specifics" of specific reading disability and specific language impairment. *Journal of Child Psychology and Psychiatry, 41,* 869–874.

McKean, C., Letts, C., & Howard, D. (2013). Developmental change is key to understanding primary language impairment: The case of phonotactic probability and nonword repetition. *Journal of Speech, Language, and Hearing Research, 56,* 1579–1594.

Melby-Lervag, M., & Lervag, A. (2012). Oral language skills moderate nonword repetition skills in children with dyslexia: A meta-analysis of the role of nonword repetition skills in dyslexia. *Scientific Studies of Reading, 16,* 1–34.

Metsala, J. L. (1997). Spoken word recognition in reading disabled children. *Journal of Educational Psychology, 89,* 159–169.

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology, 91,* 3–19.

Munson, B. (2001). Phonological pattern frequency and speech production in children and adults. *Journal of Speech, Language, and Hearing Research, 44,* 778–792.

Munson, B., Kurtz, B. A., & Windsor, J. (2005). The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without language impairments. *Journal of Speech, Language, and Hearing Research, 48,* 1033–1047.

Oostdijk, N. (2000). The spoken Dutch corpus: Overview and first evaluation. In M. Gravilidou, G. Carayannis, S. Markantonatou, S. Piperidis, & G. Stainhaouer (Eds.), *Proceedings of the Second International Conference on Language Resources and Evaluation* (Volume 2, pp. 887–893). Paris, France: European Language Resources Association.

Ramus, F., & Szenkovits, G. (2008). What phonological deficit? *The Quarterly Journal of Experimental Psychology, 61,* 129–141.

Raven, J. (2006). *Raven Standard Progressive Matrices*. Enschede, the Netherlands: Harcourt.

Rispens, J. E., & Baker, A. E. (2012). Nonword repetition: The relative contributions of phonological short-term memory and phonological representations in children with language and reading impairment. *Journal of Speech, Language, and Hearing Research, 55,* 683–694.

Rispens, J. E., & de Bree, E. (2014). Past tense productivity in Dutch children with and without SLI: The role of morphophonology and frequency. *Journal of Child Language, 41,* 200–225.

Rispens, J. E., & Parigger, E. M. (2010). Non-word repetition in Dutch-speaking children with specific language impairment with and without reading problems. *British Journal of Developmental Psychology, 28,* 177–188.

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

Schlichting, L. (2005). *Peabody Picture Vocabulary Test—III NL*. Amsterdam, the Netherlands: Harcourt.

Snowling, M. J. (2000). *Dyslexia* (2nd ed.). Oxford, United Kingdom: Blackwell.

Storkel, H. L. (2004). Methods for minimizing the confounding effects of word length in the analysis of phonotactic probability and neighborhood density. *Journal of Speech, Language, and Hearing Research, 47,* 1454–1468.

Storkel, H. L., Armbruster, J., & Hogan, T. P. (2006). Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech, Language, and Hearing Research, 49,* 1175–1192.

Thomson, J. M., Richardson, U., & Goswami, U. (2005). Phonological similarity neighborhoods and children's short-term memory: Typical development and dyslexia. *Memory & Cognition, 33,* 1210–1219.

Thorn, A. S. C., & Frankish, C. R. (2005). Long-term knowledge effects on serial recall of nonwords are not exclusively lexical. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 729–735.

van den Bos, K. P., Spelberg, H. C. L., Scheepstra, A. J. M., & de Vries, J. R. (1994). *De Klepel. Een test voor de leesvaardigheid van pseudo-woorden*. Nijmegen, the Netherlands: Berkhout Testmateriaal.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45,* 2–40.

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science, 9,* 325–329.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language, 40,* 374–408.

Walley, A. C., Michela, V. L., & Wood, D. R. (1995). The gating paradigm: Effects of presentation format on spoken word recognition by children and adults. *Perception & Psychophysics, 57,* 343–351.

Zamuner, T. S. (2006). Sensitivity to word-final phonotactics in 9- to 16-month-old infants. *Infancy, 10,* 77–95.

Appendix A

Items of the Lexical Decision Task (Word Recognition)

Pair number	Word	ND size, high/low	Frequency, high/low	Nonword	PP nonword	English
1	koffer	6, high	1.69, high	toffer	−1.32	suitcase
2	negen	10, high	1.71, high	megen	−0.99	nine
3	deken	6, high	1.62, high	beken	−1.01	blanket
4	boven	11, high	2.65, high	moven	−0.97	up
5	mager	8, high	1.68, high	bager	−1.08	lean
6	zeven	10, high	1.94, high	deven	−0.96	seven
7	kussen	7, high	1.93, high	kuffen	−1.20	pillow
8	vader	8, high	2.76, high	vazer	−1.19	father
9	keuken	8, high	1.98, high	keupen	−1.54	kitchen
10	regen	14, high	1.74, high	refen	−1.16	rain
11	hemel	8, high	2.01, high	hebel	−1.38	heaven
12	water	8, high	2.56, high	waser	−1.09	water
13	suiker	2, low	1.60, high	fuiker	−1.14	sugar
14	vogel	1, low	1.98, high	sogel	−1.25	bird
15	bodem	1, low	1.69, high	dodem	−1.45	bottom
16	nummer	2, low	1.86, high	dummer	−1.18	number
17	jongen	2, low	2.56, high	rongen	−1.41	boy
18	weinig	1, low	2.86, high	meinig	−1.53	few
19	motor	1, low	1.69, high	mopor	−1.18	motor
20	middag	1, low	1.85, high	mibbag	−1.06	afternoon
21	cijfer	0, low	1.58, high	cijger	−1.16	number
22	heuvel	3, low	1.72, high	heupel	−1.23	hill
23	midden	3, low	2.36, high	minnen	−1.19	middle
24	koning	3, low	2.00, high	koling	−1.02	king
25	gieter	6, high	0.00, low	fieter	−0.73	watering can
26	poeder	10, high	0.95, low	koeder	−1.55	powder
27	bever	12, high	0.00, low	dever	−1.23	beaver
28	veulen	7, high	0.48, low	weulen	−1.34	foal
29	wekker	8, high	0.85, low	bekker	−1.04	alarm clock
30	bobbel	7, high	0.30, low	mobbel	−1.14	bump
31	kikker	10, high	0.95, low	kigger	−1.08	frog
32	kever	9, high	0.48, low	keser	−1.12	beatle
33	kuiken	9, high	0.47, low	kuiten	−0.99	chicken
34	hagel	6, high	0.60, low	hakel	−1.17	hail
35	veter	5, high	0.70, low	veker	−1.18	lace
36	rover	9, high	0.70, low	roper	−1.41	robber
37	pudding	0, low	0.48, low	tudding	−1.26	pudding
38	puzzel	0, low	0.60, low	fuzzel	−1.38	puzzle
39	tennis	3, low	0.30, low	sennis	−1.08	tennis
40	sappig	3, low	0.70, low	tappig	−1.27	juicy
41	tijger	2, low	0.85, low	dijger	−1.51	tiger
42	kachel	1, low	1.18, low	pachel	−1.14	heating
43	haring	3, low	0.70, low	hajing	−1.16	herring
44	reiger	2, low	0.48, low	reiser	−1.22	heron
45	toeter	2, low	0.10, low	toeser	−1.15	horn
46	hoepel	3, low	0.10, low	hoekel	−1.23	hoop
47	bezem	2, low	0.60, low	bevem	−0.98	broom
48	lego	1, low	0.10, low	leko	−1.18	lego

Note. ND = neighborhood density; PP = phonotactic probability. Filler items: *koekje* (cookie), *circus* (circus), *zeeleeuw* (sea lion), *pinguin* (penguin), *varken* (pig), *zebra* (zebra).

Appendix B

Items of the Nonword Repetition Task

Orthography	IPA	Mean PP score	PP (high/low)
Two-syllable items			
1. weugof	wø:xɔf	–1.76	Low
2. kuimup	kœymʏp	–1.86	Low
3. luubuf	lybʏf	–1.84	Low
4. joefeum	jufø:m	–1.83	Low
5. feusut	føsʏt	–1.73	Low
6. hiewam	hiwɑm	–1.25	High
7. raanom	ranɔm	–1.17	High
8. geeres	xerɛs	–1.20	High
9. woosel	wosɛl	–1.18	High
10. daanes	danɛs	–1.20	High
Three-syllable items			
1. woezuunim	wuzynɪm	–1.85	Low
2. muihuuguf	mœyhyxʏf	–2.43	Low
3. soeguipem	suxœypɛm	–1.81	Low
4. nuigeusup	nœyxø:sʏp	–1.91	Low
5. veujoetup	vø:jutʏp	–1.85	Low
6. kaaroodin	karodɪn	–1.17	High
7. voopeeket	vopekɛt	–1.27	High
8. loowaamas	lowamɑs	–1.35	High
9. taanoolon	tanolɔn	–1.33	High
10. deevoenos	devunɔs	–1.29	High
Four-syllable items			
1. meufuusuinef	mø:fysœynɛf	–2.05	Low
2. suijiegoonif	sœyjixonɪf	–2.01	Low
3. juuvuigoowuf	jyvœyxowʏf	–2.13	Low
4. guiweusoegir	xœywø:suxɪr	–2.12	Low
5. fuiseuwoesut	fœysø:wusʏt	–2.00	Low
6. liejootaanig	lijotanɪx	–1.44	High
7. peewaatoopes	pewatopɛs	–1.48	High
8. liekoovoepar	likovupɑr	–1.41	High
9. saaviebeemer	savibemɛr	–1.44	High
10. kooviewaalan	koviwalɑn	–1.28	High
Five-syllable items			
1. fuugiwuinoefep	fyxɪwœynufɛp	–2.13	Low
2. soegonuifeusir	suxɔnœyfø:sɪr	–1.87	Low
3. geumuwoekuubir	xø:mʏwukybɪr	–2.14	Low
4. nuijigeufuusut	nœyjɪxø:fysʏt	–2.01	Low
5. jeunimeusuifir	jø:nɪmø:sœyfɪr	–1.94	Low
6. baamerienooves	bamɛrinovɛs	–1.34	High
7. geerutievaanot	xerʏtivanɔt	–1.41	High
8. tieloniedaanag	tilɔnidanɑx	–1.31	High
9. wookaloemoodon	wokalumodɔn	–1.34	High
10. beemonievoekes	bemɔnivukɛs	–1.38	High

Note. IPA = International Phonetic Alphabet; PP = phonotactic probability.