LEXICAL REPRESENTATIONS AND DEVELOPMENT: 
THE EMERGENCE OF RIME PROCESSING

U. Goswami and B. De Cara 
Institute of Child Health, University College London, U.K.

ABSTRACT
This study investigated the relationship between the emergence of phonological awareness in young children and the nature of inter-item phonological similarity relations in the English monosyllabic lexicon. In most discussions of phonological similarity relations, rime neighbours like *pot*, onset-vowel or lead neighbours like *cough*, and consonant neighbours like *kit*, are all considered equal neighbours of a target word like *cot*. Using two different metrics for calculating similarity neighbourhoods, it was established that rime neighbours predominate in the English phonological lexicon for monosyllables. We then examined the effects of two factors on the development of rime awareness in young children; phonological neighbourhood density (language specific) and sonority profile (language universal). Our experiment crossed sonority profile (good, poor) with neighbourhood density (dense, sparse). Five- and 6-year-old children were tested. Significant effects of both neighbourhood density and sonority profile on the development of rime-level processing were found.

1. INTRODUCTION
The child’s awareness of the phonology of his or her language is one of the most important predictors of that child’s progress in learning to read and to spell, but cross-linguistic research is increasingly showing that the phonological units that are highlighted by different languages may vary [1]. Phonological awareness is usually measured by tasks that require a child to reflect upon or to manipulate the component sounds of spoken words. Studies in English suggest that there is a developmental progression from phonological awareness of ‘large’ segments or units of phonology (syllables, onsets and rimes), to phonological awareness of ‘small’ segments or units (phonemes) [2]. Given its importance for literacy, surprisingly little work has been done on the linguistic and lexical factors that might determine the development of phonological awareness in all children in different languages.

One proposal is that phonological awareness may emerge as a result of ‘lexical restructuring’ processes that are part of language acquisition [3]. According to ‘lexical restructuring theory’ (LRT), segmental phonology is represented at an increasingly fine-grained level with development. LRT argues that children’s first words represent global phonological characteristics, but that as more and more words are acquired, children begin to represent smaller segments in words such as syllables and phonemes. Although the LRT account is inconsistent with recent studies of linguistic processing in infants [4], it is highly consistent with the literature on the development of phonological awareness.

1.1. Phonological Neighbourhoods and English Monosyllables
As LRT proposes that implicit comparisons between similar-sounding words constitute the basis for the emergence of phonological awareness, it seems logical that the nature of the phonological neighbours in the child’s lexicon in different languages will influence this developmental process. The traditional similarity metric for defining a phonological neighbourhood considers neighbours to be words that differ by the addition, deletion or substitution of a single phoneme [5]. According to this ‘N+/-1’ metric, rime neighbours like *pot*, onset-vowel or ‘lead’ neighbours like *cough*, and ‘consonant’ neighbours like *kit*, are all considered to be equal neighbours of a target word like *cot*. However, an alternative definition can be proposed based on a linguistic analysis that codes monosyllables in terms of the phonological units onset, nucleus, and coda [6]. The latter similarity metric is called here ‘Nvec’. The chief difference psychologically would be that, whereas words like *spot* and *trot* would count as rime neighbours of *cot* in the Nvec database, they would not count as rime neighbours of *cot* in the N+/-1 database [7].

To examine the characteristics of English phonological neighbours for monosyllables, we recently analysed the corpus of single-syllable words in the Luce & Pisoni [8] database of spoken English forms in terms of rime neighbours (RN), onset-vowel neighbours (OVN), and consonant neighbours (CN) in dense versus sparse neighbourhoods respectively. We used two measures of phonological neighbourhood, N+/-1 and Nvec. The results were broadly similar (see Table 1). Rime neighbours predominated in dense neighbourhoods according to both similarity metrics (although these analyses are by type, analyses by token produced almost identical results [7]). Thus developmentally, words in dense neighbourhoods might experience more pressure for early lexical restructuring to the rime level than words in sparse neighbourhoods. Although the similarity indices shown in Table 1 are derived from an adult lexical database, we have also calculated the number of rime, onset-vowel and consonant neighbours for two different estimates of early-acquired words [8,9]. The results in terms of percentage of each type of neighbour are virtually identical to those presented in Table 1 [10].
Table 1: Phonological neighbourhood in English monosyllabic words (Nvcc and N+/−-metrics)

<table>
<thead>
<tr>
<th></th>
<th>All monosyllabic words (n = 3072)</th>
<th>Dense N (Nvcc ≥ 32) (n = 570)</th>
<th>Sparse N (1 ≤ Nvcc ≤ 13) (n = 619)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nevc</td>
<td>RN</td>
<td>CN</td>
</tr>
<tr>
<td>M</td>
<td>22.5</td>
<td>12.2</td>
<td>3.8</td>
</tr>
<tr>
<td>SD</td>
<td>10.1</td>
<td>7.8</td>
<td>3.2</td>
</tr>
<tr>
<td>%</td>
<td>100</td>
<td>51.4</td>
<td>16.4</td>
</tr>
<tr>
<td>N=1</td>
<td>13.4</td>
<td>5.7</td>
<td>3.8</td>
</tr>
<tr>
<td>SD</td>
<td>8.7</td>
<td>4.6</td>
<td>3.2</td>
</tr>
<tr>
<td>%</td>
<td>100</td>
<td>41.2</td>
<td>27.4</td>
</tr>
</tbody>
</table>

1.2. Phonological neighbourhoods and the development of phonological awareness.

If the statistical patterns demonstrated in Table 1 actually affect the development of phonological awareness, then children should find it easier to decide that words in dense neighbourhoods rhyme than to decide that words in sparse neighbourhoods rhyme. We used two different phonological awareness tasks to test this hypothesis. One was the oddity task [11], in which children must select the ‘odd word out’ in terms of the rhyme from a triple of words (e.g., pit, hit, got). The second task was the same-different judgement task [12], in which children must decide whether two spoken words share a target sound or not. Both tasks were chosen because are meant to tap ‘epilinguistic’ processing [13], which theoretically results primarily from spoken vocabulary growth and associated changes in the familiarity of individual lexical items and inter-Item phonological similarity relations. We also varied the sonority profile of the syllable in our experiment, considered a linguistic universal. We created stimuli with rimes that had either a ‘good’ or a ‘poor’ sonority profile, contrasting rimes ending in liquids with rimes ending in stops.

2. Method

2.1. Participants

A group of 45 children took part in the study. The mean age of the younger group was 5 years 4 months (13 girls, 8 boys, s.d. 4 months), and the mean age of the older group was 6 years 6 months (13 girls, 11 boys, s.d. 4 months). Performance on a standardised vocabulary test (British Picture Vocabulary Scales, mean = 100, s.d. = 15) was 5-year-olds: 109.5, s.d. 14.6; 6-year-olds: 108.0, s.d. 15.1. The children were unselected, comprising all those enrolled in the participating school whose parents returned a consent form allowing them to take part in the study. Testing was conducted during the second half of the summer term.

2.2. Procedure

Each child participated in 2 short testing sessions. In the first session the oddity task was administered. In the second session the same-different judgement task was administered. A period of at least a week separated the two sessions. The words for each trial were recorded by a native female speaker of British English and then digitised for computer presentation using Cool Edit™ 96 (Syntrillium Software Corporation). Before each trial, the children saw a row of asterisks in the centre of the computer screen, which disappeared when the trial began. The stimuli were presented through headphones. For each oddity trial, the child had to press the space bar and say the word when they knew the answer. For each same-different judgement trial, the child had to press the space bar and say ‘yes’ or ‘no’ when they knew the answer.

2.3. Tasks

2.3.1. Oddity Task

This task was based on triples of words. Each triple was spoken by the computer, and the child was asked to select the odd word out. Children were told that the odd word would not rhyme with the others. There were 36 experimental trials overall, 9 for each of the sonority profile x neighbourhood density (SP x ND) manipulations. Within each SP x ND category, we varied whether the triples were based on a vowel change (e.g., pit, hit, got), on a codata change (e.g., meat, weak, seat), or both a vowel and codata change (e.g., peak, dot, not). Trials were not blocked by SP x ND category, but varied in a semi-random order which also varied the position of the odd word systematically across the experiment. Detailed feedback was provided in 5 training trials given prior to the experimental trials. No feedback was given on the experimental trials.

2.3.2. Same-Different Judgement Task

This task was based on pairs of words. Each pair was spoken by the computer, and the child was asked to judge whether the two words shared a final sound or not. Children were told whether the final sound was a rhyme or a codata (for the codata, they were told to listen for “just the same sound at the end”). There were 80 experimental trials overall, 10 ‘yes’ and 10 ‘no’ trials for each of the SP x ND manipulations. Trials were not blocked by SP x ND category, but varied in a semi-random order which also varied whether a ‘yes’ or ‘no’ judgement was required. Within the session, rhyme trials (N = 40) and codata trials (N = 40) were blocked and counter-balanced.
for order of presentation across children. Detailed feedback was first provided in 6 training trials. No feedback was given on the experimental trials.

2.4. Stimuli

The same pool of CVC words was used as a basis for both of the tasks. Only words that were highly familiar to young children were selected. In each task, triples (oddball task) or pairs (same-different judgement task) of words either had a good or poor sonority profile, and were either from dense or sparse neighbourhoods. The words were selected from a large auditory database (Luce & Pisoni, 1998) containing 3619 monosyllabic words. Monosyllables with either no onset or with a complex onset or coda (i.e., with a CCC structure) were excluded because they were relatively rare, and because the sonority of CCC structures is unusual. This left 3072 monosyllables (85% of total) with the following distribution; CVC 44.3%, CCVC 24.4%, CCVC 15.7%, CCVCC 6%, CV 5.9%, CV 3.8%. Stimuli were selected for the experiments if they ended with either a liquid ('good' SP) or a stop ('poor' SP). To vary rime neighbourhood density, we initially aimed to select stimuli from neighbourhoods that were either 1 s.d. above or below the mean of 12 rime neighbours (s.d. = 8). In practice, this did not yield sufficient stimuli for either the sparse or dense N comparisons. This was because most words with RN < 5 are unfamiliar to young children (e.g., daub, lour, moll), while few words with RN > 20 end in liquids. We therefore selected the best contrast possible in view of our other constraints, yielding a mean RN for dense stimuli of 21.5 and a mean RN for sparse stimuli of 8. The number of vowels used was limited for each SP x ND category in order to avoid large disparities in vowel format between categories, and the same vowels were used as far as possible. Vowel quality (short, long, diphthong) and consonantal features (manner, place of articulation, voicing) were matched between target and distractor words across stimuli.

3. RESULTS

Results are discussed by task in terms of number of errors and mean reaction times. The few occasions on which children requested to hear the stimuli again were roughly equivalent across SP x ND category, and were counted as errors. Reaction times that were more than 2 s.d. from the mean were excluded from the analyses.

3.1. Oddity Task

The mean number of errors in the oddity task by SP x ND category for the 5- and 6-year-olds respectively are shown in Table 2. The mean RTs by SP x ND category for the 5- and 6-year-olds respectively are shown in Table 3. Two 2 x 2 x 2((Age, 5 years, 6 years) x SP (good, poor) x ND (dense, sparse)) Anovas were run, taking the mean number of errors and the mean RT as the dependent variables, respectively. The error analysis showed main effects of age, $F(1,43) = 8.19$, $p < .01$, sonority profile, $F(1,43) = 6.16$, $p < .05$, and neighbourhood density, $F(1,43) = 5.34$, $p < .05$. There were no interactions. Younger children made more errors than older children (3.38 versus 2.17), good SP items were processed more easily than poor SP items (2.57 errors versus 2.98), and words in dense neighbourhoods were processed more easily than words in sparse neighbourhoods (2.58 errors versus 2.97). The RT analysis showed a main effect of age only, $F(1,43) = 7.52$, $p < .01$. The older children were faster than the younger children (1098 ms versus 1412).

3.2. Same-Different Judgement Task

Performance in the same-different judgement task was at ceiling for the rime version of the task at both ages. Performance in the coda version of the task was at chance for the younger children. This finding is consistent with the many demonstrations that phoneme judgements are very difficult for young or pre-reading children. The RT data were unreliable for the same-different judgement task as the task was either too easy (rime task) or led to a high proportion of errors and therefore yielded rather little RT data (coda level). Coda-level performance by the older children only showed that significantly more errors were made for
good SP stimuli than for poor SP stimuli. There were no apparent effects of neighbourhood density. There were no other significant effects.

4. DISCUSSION

In our study, children were significantly more accurate at making judgements about rhyme for words from dense neighbourhoods than for words from sparse neighbourhoods in the oddity task. Overall, therefore, neighbourhood density effects can emerge in simple phonological awareness tasks. The RT data was too noisy to enable comparison of judgement times for stimuli from dense versus sparse neighbourhoods. In the adult speech processing literature, there is typically a speed cost for processing stimuli from dense neighbourhoods. Data currently being collected in our laboratory suggest that dyslexic children (mean age 9 years 6 months) show a similar speed cost in the oddity task, although accuracy differences have disappeared by this age. However, these findings must be considered preliminary as control data from normal readers is still being collected. Regarding sonority profile, rhymes with good SP were processed more accurately than rhymes with poor SP in the oddity task. However, rhymes with poor SP were processed more accurately than rhymes with good SP in the same-different judgement task. One possible explanation may be that the same-different judgement task and the oddity task make different processing demands on the auditory versus articulatory systems. The oddity task may tap auditory processing, while the same-different judgement task may be more easily solved by articulatory processing. Interestingly, the children took almost twice as long to make a judgement in the coda version of the same-different judgement task compared to the coda change trials in the rime oddity task. Our current work is continuing to investigate the effects of phonological neighbourhood density and sonority profile on the development of performance in a variety of different phonological awareness tasks, and we are also examining children’s reading and spelling of monosyllables when these factors are varied.

5. ACKNOWLEDGEMENTS

We thank Belmont School, London, for participating in our study, and Andy Faulkner and Jill House for their help with stimulus preparation. Support for this research was partly provided by a Fyssen Foundation award to Bruno De Cara, Corresponding author: Usha Goswami, Institute of Child Health, University College London, 30 Guilford St., London WC1N 1EH. Tel. 44 171 831 0975. Fax 44 171 831 0975, e-mail: u.goswami@ich.ucl.ac.uk.

6. REFERENCES