JSLHR Papers in Press. Published August 26, 2010, as doi: 10.1044/1092-4388(2010/08-0148)
The final version is at http://jslhr.asha.org.

Running head: SPOKEN WORD CLASSIFICATION

Spoken Word Classification in Children and Adults.

Julia M. Carroll

Joanne M. Myers

University of Warwick, UK

Address for Correspondence:

Julia M. Carroll

Department of Psychology

University of Warwick

Coventry

CV4 7AL

UK

Email: **J.M.Carroll@warwick.ac.uk**

Telephone: (+44) 2476 523 613

Fax: (+44) 2476 524 225

Abstract

Purpose: Preschool children often have difficulties in word classification, despite good speech perception and production. Some researchers suggest they represent words using phonetic features rather than phonemes. We examine whether there is a progression from feature based to phoneme based processing across age groups, and whether responses are consistent across tasks and stimuli.

Method: In Study 1, 120 3 to 5 year old children completed three tasks assessing use of phonetic features in classification, with an additional 58 older children completing one of the three tasks. In Study 2, all of the children, together with an additional adult sample, completed a nonword learning task.

Results: In all four tasks, children classified words sharing phonemes as similar. In addition, children regarded words as similar if they shared manner of articulation, particularly word-finally. Adults also showed this sensitivity to manner, but across the tasks there was a pattern of increasing use of phonemic information with age.

Conclusions: Children tend to classify words as similar if they share phonemes or share manner of articulation word finally. Use of phonemic information becomes more common with age. These findings are in line with the theory that phonological representations become more detailed in the preschool years.

Spoken Word Classification in Children and Adults

There is a body of evidence that children tend to classify words sharing phonetic features as similar (Snowling, Hulme, Smith, and Thomas, 1994). Some researchers argue that this reflects a fundamental difference in phonological representations of young children in comparison to those of adults (Storkel, 2002), while others argue that phonological representations are adult-like from the second year of life (Bailey & Plunkett, 2002; Swingley, 2009). This paper investigates whether children's phonological classification responses are consistent across different ages, tasks and stimuli. Predictable changes with age that are consistent across a range of tasks and stimuli would strengthen the argument that phonological representations change in the preschool years.

Skills at different ages: Phonological skills in infancy and early childhood

Speech Perception and Production in Infancy. There is a clear contrast between the good speech perception abilities shown by infants and the difficulties in sound segmentation and classification shown by three and four year old children. Even 4 month old children appear to perceive phonemes categorically, (Eimas, Siqueland, Jusczyk, & Vigorito, 1971). However, some studies have suggested that infants do not always process detailed phonetic information when learning new words. For example, 14 month old infants do not seem to detect a mispronunciation of a newly learnt word, while 8 month old infants (who are not yet linking words and objects) do detect the change, apparently showing greater sensitivity than the older children (Stager & Werker, 1997). In terms of early speech production, first words tend to follow a limited set of phonological patterns or 'frames' which are often quite different from adult realisations (MacNeilage & Davis, 2000; Piske, 1997) and often demonstrate high intra-word variability (Sosa & Stoel-Gammon, 2006). These factors suggest syllabic, rather than phonemic, representation for early words.

Other research suggests that children do represent phonetic detail from their second year of life. For example, in a preferential looking task, 19 month old infants are less likely to look towards a target when the target word is mispronounced, even by a single phonetic feature (White & Morgan, 1998). It may be that in the mispronunciation detection research mentioned above, the added cognitive load of word learning causes 14 month old infants to focus less on phonetic detail (Fennell & Werker, 2003). Many researchers conclude that, even if children do show initial holistic processing, this is overcome by the end of the third year at the latest, and phonological representations are then adult-like (e.g. Gerken, Murphy & Aslin, 1995).

In contrast, Walley's (1993) Lexical Restructuring Hypothesis argues that children gradually increase detail in phonological representations throughout the preschool years, and that this process is accelerated by vocabulary growth and literacy instruction. Holistic representations would be present in some form, for some words, throughout the pre-school and early school years.

The Preschool Years. Despite good speech perception and production, there are reasons to believe that three year old children's representations of words are not adult-like. Three to five year old children show substantial changes in the way they approach phonological awareness tasks. Three and four year old children find it difficult to say whether words share a given phoneme or segment (such as a rime). One study shows only 28.9% of children at 3;8 years were above chance on a rime matching task, and only 5% of the group were above chance on an initial sound matching task, despite a 50% chance rate on both tasks and extensive feedback and memory support (Carroll and Snowling, 2001).

There are at least two possible reasons for these difficulties. First, phonological awareness tasks require some degree of reflection on words, though perhaps relatively non-conscious or 'epilinguistic' (e.g. Moraias, 2003; Morton and Frith, 1993). It may be that three and four year old children find it difficult to reflect upon words out of context. Alternatively, it may be that the detection of similarities between words places greater demands on phonological representations and processing than producing or perceiving individual words. In order for the same phoneme to be recognised in two different words, children would have to have stored the information about that phoneme in the same way in the two words and, to solve these tasks successfully, words would have to be represented in a systematic way (e.g. as a series of phonemes or segments) rather than an idiosyncratic, word by word way. It may be preschoolers store words in terms of sets of phonetic features. Features are potentially less consistent across words because of changes in the realisation of particular features in different articulatory contexts.

If children move from idiosyncratic, feature-based representations to more systematic, phoneme-based representations in the late pre-school and early school years, one would expect to see specific changes to the way children across this age range approach classification tasks. For instance, younger children would tend to rate as similar two words sharing several features, but not necessarily an entire phoneme, while older children would rate the words sharing phonemes as more similar. It is therefore useful to examine different age groups on the same tasks. It could also be useful to have an adult comparison group to be certain that adults would use phonemic strategies on these tasks.

The Role of Task: Previous research examining feature-based classification

As noted above, studies examining phonological skills in infants and preschoolers vary not only in the age of the participants, but in the types of tasks used. Preschoolers are typically asked to compare or classify words, whereas younger children are tested on tasks such as mispronunciation detection. It is therefore worth considering findings in light of the different types of tasks used.

In a preferential looking task with target words mispronounced by one, two or three phonetic features, 19 month old infants were sensitive to the number of shared features in mispronounced words (White and Morgan, 2008). They looked at the target object more when words were only one feature different from the target than when the words differed by two phonetic features, and looked least of all when the word differed by three features. Similar results have been found by other researchers (e.g. Halle and de Boysson-Bardies, 1994; Mani and Plunkett, 2007; Swingley, 2005; Vihman, Nakai, DePaolis and Halle, 2004). However, the paradigms used by these researchers focus on *specificity within* phonological representations, rather than *similarities across* phonological representations. Very few studies examine whether infants perceive different words as sounding similar.

In a list listening task with nine month olds, infants listened longer to sequences of words that shared the same manner of articulation in the initial phoneme than to rhyming sequences or those that had no phonological relationship (Jusczyk, Goodman, and Bauman, 1999). In a further study, there was no difference in listening time between sequences sharing initial phonemes and sequences sharing only manner word-initially. Overall, infants showed sensitivity to shared manner across words. This task has similarities with the phonological matching and judgment tasks used to assess phonological awareness, in that it relies on perceived sound similarities across words.

Older children can perceive similarities between words that do not share phonemes. For example, many five year old children who passed a standard alliteration detection task were not able to consistently discount distracter words that did not share a phoneme, but were globally phonologically similar (Byrne and Fielding-Barnsley, 1993). These results were replicated with rhyming words (Cardoso-Martins, 1994; Carroll and Snowling, 2001). Global similarity, in this context, is based upon the acoustic perceptual similarity ratings of Singh and colleagues (Singh and Woods, 1971; Singh, Woods, and Becker, 1972), but shared phonetic features provide a good predictor of phonological similarity ratings in adult samples (Bailey and Hahn, 2005).

Storkel (2002) examined the classification of words by 20 children with an average age of 4;8 years. Children were given a word and asked to say whether it sounded like a target word or not. Children had a tendency to classify words as similar if they shared manner of articulation in the coda. However, this occurred only in words from sparse phonological neighbourhoods. The author concluded that manner of articulation is the first feature which is consistently detected across words, but other features and information combine with this as children grow older and their lexicons become more densely populated, though this hypothesis remains to be tested on older children.

These classification tasks involve comparison between different words, but they remain limited measures of children's understanding of sound similarities, because the words are provided for the children and a 'yes or no' response is requested. Other approaches examining which words children regard as sounding similar could include a production task, where the child is asked to provide a similar sounding word themselves, and a memory task in which confusions between different pair of words are analysed. Tasks of these types are included in the present study. To our knowledge, word production tasks in children have not been analysed with regard to shared phonetic features, but this task avoids imposing particular phonological pairings on the child – their response is open ended, and thus it provides a useful measure of phonological processes.

A nonword learning task provides a potentially sensitive way of assessing the development of new phonological representations. For example, in a nonword learning task with adults, Magnuson, Tanenhaus, Aslin and Dahan (2003) showed a decrease in eye movements to rhyming distractor items over training, while eye movements to distractor items starting with the same initial sound actually increased with training, indicating an increased focus on initial sounds as the words became familiar. Similarly, young children learn nonwords with high phonotactic probabilities more quickly than those with low phonotactic probabilities (Storkel, 2001). Examining the errors made during nonword learning would help to provide evidence of which words are represented as similar without requiring an explicit judgment from the child.

Previous research analysing memory errors in a nonword learning task provides some evidence for the hypothesis that classification changes with age: Treiman and Breaux (1982) found that young children tend to confuse 'globally' similar sounding nonwords, while adults confuse nonwords that share initial phonemes, though this study did not directly examine the role of phonetic features. This task avoids explicit classification and hence reduces the extent to which results depend on the child's view of what the experimenter wants to hear.

Word Initial and Word Final Consonants

In addition to variation according to the task used, there is also evidence of variation according to the position of a consonant within a word. For example, Storkel (2002) provides evidence that children process word-initial and word-final sounds differently. Ziegler and Goswami (2005) develop Walley's (1993) Lexical Restructuring Hypothesis by suggesting that children show an intermediate stage between fully holistic representations and phonemic representations in which they represent words in term of the onset (initial consonants) and rime (vowel and any following consonants). In this theory, children show awareness of syllables, then onsets and rimes, and finally phonemes. For example, when selecting an 'odd word out' from a group of three, children found it easier to correctly select a word with differing word-final consonant if the vowel of the word also differed (Kirtley, Bryant, MacLean, and Bradley, 1989). However, the vowel made no difference to the difficulty level of the word-initial oddity task. This was taken as evidence that children represent the word onset separately from the vowel and following consonants, which are thought to be represented as a single unit (known as the rime).

There is certainly evidence to suggest that codas could be represented in less detail than onsets. Many languages do not use codas at all, and there are no languages in which syllables must have a coda. Final consonant deletion is a common feature of typically developing children's speech until the age of three (Grunwell, 1997). Additionally, final consonant devoicing occurs in many languages (Locke, 1983), and in some languages, word-final consonants have disappeared over time (Hock, 1975, cited in Locke, 1983), or have been partially replaced by glottal stops. For example, Shockey and Bond (1980) found that British mothers replaced word-final [t] with a glottal stop 44% of the time while talking to their children. In line with these indications, adults are more likely to confuse newly learnt words that differ word finally than word initially, indicating less focus on word final consonants (Creel & Dahan, 2010). In contrast to these findings, however, eye tracking research demonstrates that 18 month old children and adults have high sensitivity to both word-initial and word-final mispronunciations (Swingley, 2009). The different tasks used in these experiments may provide a potential explanation for these contrasting results; one assesses detection of a consonant within a well-known word, while the other is more demanding, requiring an explicit decision on a newly learnt word.

If codas are represented less consistently across words than onsets, children may be less able to use them reliably in matching tasks. It may be that the addition of a vowel provides the extra information necessary to make an accurate matching decision. This is less important for word onsets because they are more consistently (or phonemically) represented. This difference between word-initial and word-final sounds could provide an explanation for Kirtley et al.'s (1989) finding that children find it easier to match words on the basis of their rime rather than their coda alone, but do not show this pattern for onsets. Since the mispronunciations used in Swingley's (2009) eye tracking experiment all shared a vowel with the target word, this could also provide an explanation for sensitivity to codas that was demonstrated.

Summary

Previous research has demonstrated the following: three and four year old children find phonological classification tasks relatively difficult, despite good speech production and perception skills; young children are sensitive to sub-phonemic similarities and use these in phonological awareness tasks; and children may represent codas in less detail than onsets. However, most of these findings are based upon studies in which single age groups are tested on one particular task. It is not clear the extent to which these patterns in young children may represent task related demands, such as short-term memory, rather than real differences between adult and child phonological representations.

In order to address this question, we examine performance of different age groups of children on a range of different perception and production tasks, to reduce the possible influence of task demands. We also examine word initial classification and word final classification separately. Finally, in Study 2, we include a task based on the work of Treiman and Breaux (1982) looking at word learning processes rather than similarity judgments. This study also provides an adult comparison group.

Study 1

This study investigates how children respond to shared articulatory features in phonological classification tasks. It is anticipated, as shown in previous research, that children will show a tendency to classify words on the basis of shared manner. It is also anticipated that this will be particularly true word-finally, perhaps because word final consonants are represented in less detail. Vowels may also influence word final classification more than word initial classification; this will be assessed in the Word Production task and the Sound Families task.

If phonological representations do indeed change with age, one would expect that the younger children would be more likely to classify words on the basis of phonetic features, while the older children would be more likely to rely on phonemic information alone. Three age groups of children were included: pre-readers, beginning readers, and children who had received at least a year's literacy instruction (though time constraints meant that the latter group completed only one of the three tasks). If phonological representations vary with age or literacy experience, one would anticipate an interaction between classification of the different items and age group. The groups would also differ in terms of the types of words produced in the word production task, with older children producing more words that share full phonemes while younger children produce words sharing phonetic features.

All of the tasks require comparison between different words, but they vary in task demands: in the Forced Choice task, two individual words are compared, while in the Sound Families task, a test word is compared to a group of words (which may make phonological similarities clearer). The Word Production task asks a child to produce a similar sounding word themselves, thereby tapping production skills. If similar patterns are shown across the three tasks, this provides support for responses being motivated by underlying representations rather than extraneous task demands.

Method

Participants

One hundred and seventy-eight children from state-run primary schools and nurseries in Warwickshire, UK took part in the study. There were 62 children in nursery and 58 children in reception year, the first year of compulsory schooling in the UK. 58 children from Year 1 took part in the Forced Choice task. Thirty-five of the nursery children, 25 of the reception class children and 29 of the Year 1 children were male. The mean age of the nursery children was 3;6 years, with a range from 3;5 years to 4;9 years. The mean age of the reception class children was 4;6 years, with a range from 4;3 years to 5;7 years. The Year 1 had an average age of 5;8 years, with a range from 5;1 years to 6;4 years.

Children with known hearing difficulties (as reported by teachers or parents) were excluded from the sample. Children with English as a second language (n = 2) and bilingual children (n = 3) were included if their score on an expressive vocabulary test was within the average or above average range. These children were all in the reception class age group, but analysis of the group showed no differences between these children and the others in the same age group, and so their data is retained.

Procedure

Children were tested individually in a separate room or quiet corner. Most of the testing was carried out by the second author, who had a degree in psychology but no specialist training in speech and language development. The tasks were given in twelve sessions of around ten minutes each over a period of two to three weeks, and children were given sticker rewards.

Tasks

Children were asked to complete eight tasks in total, though only seven are reported here in detail[1]. The first task was an unscored practice task used to introduce the concepts of 'same' and 'a little bit the same'. Two further tasks were considered background measures of language and literacy level (Picture Naming and Letter Knowledge). The three experimental tasks were two measures of word classification (Forced Choice Classification and Sound Families) and a measure of phonetic effects in word selection and production (Word Production). They are presented in Table 1 for reference. A fourth experimental task, Memory Confusions, is also included in this table and is reported in Study 2.

Background measures. The children initially completed the Picture Naming task from the Wechsler Preschool and Primary Scale II (Wechsler, Rust, and Golombok, 2003), as a measure of expressive vocabulary. They were also assessed on their letter knowledge as an index of emergent literacy knowledge. The children saw the 26 letters and were asked to say what they were. Both letter names and sounds were acceptable answers, and children were encouraged to give both.

Introduction task. This task introduced key terminology for the experimental tasks, specifically "same", "different" and "a little bit the same". Three pairs of coloured cubes were used. Two of the pairs were similar in colour (e.g. light blue and dark blue). Children were asked to point out a cube that was the same as another cube. They had to show which cube was 'a little bit the same' as the dark blue cube (i.e. the light blue cube) [2].

Word Production. An open-ended word production task was used to assess which words children would regard as similar sounding if they were given free choice. Four CVC English words (bed, shop, nut and that) were chosen as cue words. These words were selected as words that would be known to three year old children that included a variety of different types of phonemes. 'That' was included as a common word with a relatively uncommon initial phoneme. The words all appeared in a database of speech directed at children under four years old (Carroll and Vousden, in preparation), with varying frequencies ('bed': 561, 'nut': 25, 'shop': 281 and 'that': 17802 occurrences per million words).

Children were asked to say words that 'sounded like' the cue word. They were given 45 seconds for each cue word. The children's answers were recorded and later transcribed. Words were placed into three broad categories: phonologically related, semantically related or unrelated. The phonologically related words were examined to see which phonological relationships were most commonly used by children.

Forced Choice. The task is modelled on the one used by Storkel (2002). Children are shown a character and told that she or he likes words that sound 'a little bit like' the target word. They then hear several words that vary in their phonological relationship to the target word. Children are asked to place similar words with the character, and dissimilar words in a trash can. Our version of the task was presented on a computer screen, with pictures representing the character, trash can and test words.

Storkel's task was modified in two ways. First, 'unrelated' words (that shared no consonant features, only a vowel sound) were included in the stimulus set to allow an estimation of base rates of 'sounds similar' responses to words that share only a vowel. This meant that for each target word set there would be five trials, presented in random order: identical word, shared phoneme, shared manner, shared place or unrelated word. In line with the original study, place of articulation was defined broadly, in terms of labial, coronal or dorsal consonants. The vowels always matched that of the target item. The items for this task can be seen in Appendix 1. Given the other constraints, it was not possible to control voicing systematically when creating the items. In a further change from the original task, two parallel versions of the task were created to allow assessment of a wider range of target words. Nursery and Reception children were randomly assigned to complete either version A or version B. All of the Year 1 children completed version A. In line with the design of the Storkel study, two of the target words in each version were from sparse phonological neighbourhoods and two were from dense phonological neighbourhoods, based on ratings from the CELEX database. The children were asked to complete the task in two sessions, one for the shared body words and one for the shared rime, with order counterbalanced. There were ten training and practice items presented, with feedback, before the first session.

To ensure that the words were not completely new, they were used within two stories that were read aloud to the children. These stories included the two nonwords used in the practice items (vich and biff) as character names. This procedure follows that of Storkel (2002).

The stimuli words were recorded in a soundproof room and pronounced by colleagues with standard Southern English accents, free of any strong regional dialect. The target items and the test items were recorded by different individuals. The task was presented on a laptop computer, with stimuli presented through headphones, with volume adjusted for each child. Instructions and feedback were given orally by the researcher.

Children were excluded from the data analysis if they showed a lack of understanding of the task requirements. This could be demonstrated either by a general lack of response to task instruction or by not rating the same word as sounding the same consistently.

Sound Families. This is the second classification task. It followed a similar pattern to the Forced Choice task, except that the children were asked to match words to a target group, or 'family' of words sharing phonological material. This makes the shared phonological material more explicit for the child. Each ‘family’ contained a set of words sharing either the same body or the same rime. These were presented together on a computer screen, with the sounds accompanied by five identical cartoon animals. The sounds were played one after the other to emphasise their similarity. A series of six words was then presented one at a time and the children had to decide whether they belonged in the group or not. Three of these words belonged in the group as they shared the same body or rime and were regarded as control items. The other three “test” words varied in their phonological relationship with the group. In the body condition (for example *bead, beak, bean, beam* and *beach*) these words contained onsets that either share manner or place of articulation or were unrelated. Half of these words had the same vowel and half had a different vowel. The stimuli are shown in Appendix 2. In the rime condition the pattern was the same but the match was made on the coda.

Children were randomly assigned to the body or rime condition. Within each condition, there were two sessions. Each session began with a practice set, during which feedback was given. Children had to score five out of eight items correct on the practice set to proceed to the test sets.

The stimulus words were recorded in a soundproof room and pronounced by individuals with standard southern English accents. The children could listen to the words as many times as they chose.

Results and Discussion

Background measures

Both age groups showed an average to slightly above average score on the Picture Naming task (Nursery mean scale score: 12.41 (SD: 2.33); Reception mean scale score: 11.33 (SD: 2.79), with 10 being average). As anticipated, these two groups showed clear differences in their emergent literacy knowledge. The nursery children showed low levels of letter knowledge, with 22 of the 62 children tested being unable to recall more than one letter. The reception children showed good letter knowledge, with 31 of the 58 children knowing at least 24 of the 26 letters. The median numbers of letters known were 4 for the nursery children and 24 for the reception class children.

Word Production

This task provides an opportunity to examine which types of words are selected as sounding similar when children have an open choice. Sixty-nine of the 120 participants produced at least one word in response to the cue words in the Word Production task. Many of the nursery class children showed a lack of understanding of the task, with only 25 of the 62 children producing a word. Most of the reception class children were, however, able to complete the task, with 43 out of 58 producing at least one word. In total 919 words were produced by these 69 children, with the median number of words produced being five. The number of words produced by individual children ranged from 0 to 24, with a standard deviation of 7.15, and 42 children producing more than 10 words in total. There was therefore a wide range, and some evidence of a bimodal distribution, with most participants producing either no words or several words. If only the children who produced at least one word were included in the analysis, the mean number of words produced was 11.75, and the median was 12.

Of the words produced, 849 (92.4%) had a phonological relationship with the cue word. Two percent had a semantic relationship, but no phonological relationship, to the cue word (e.g. shop – pay). The remaining 5.7% had no phonological relationship or semantic relationship with the cue word (e.g. nut – radio). Of the words with a phonological relationship to the cue word, 709 (83.8%) were CVC words. The analysis concentrated on these words, the characteristics of which are shown in Table 2[3]. The majority (441, or 62.2%) of these words were words or nonsense words that rhymed with the cue words. Rhyming was therefore the primary relationship that was considered as ‘sounding similar’ to the cue word.

One hundred and thirty-four words shared a vowel with the cue word but did not rhyme. Most of these (106, or 18.4% of words sharing a vowel) shared manner of articulation in the coda. There was no clear pattern of relationship in the onset consonant in words sharing vowel sounds.

One hundred and thirty-four words (18.8% of CVC words produced) had a different vowel from the cue word. The majority of these words (82, or 61.2% of different vowel words) shared an onset with the cue word. The words produced as similar sounding words by children of this age therefore predominantly rhyme with the cue word. Other substantial word types are words sharing vowel and manner of articulation in the coda and words sharing an initial consonant.

Of course, the frequency of words produced may depend on the natural characteristics of the language and the availability of different types of words to be produced. There are two ways to overcome this problem. The first is to examine only nonsense word responses, and the second is to provide an estimate of the types of similar words available in young children's lexicons. These analyses are presented below.

Arguably, nonword responses are the purest measure in this task of children's sound similarity judgments, as within this situation a child is creating a word specifically to sound similar to the cue. Two hundred and fifty-one nonsense words were produced, with 209 of them being CVC syllables. One hundred and eighty-two of these words (87.1%) shared a vowel with the cue word, and 138 (66.0%) rhymed with the cue word. Of the 44 remaining words that shared a vowel, 37 of them (or 17.7% of the words sharing vowels) shared manner of articulation word finally. These rates are very similar to the rates for the overall responses, suggesting that this pattern is not due to language biases. In both cases, children are likely to produce rhyming words when asked to consider similar sounding words. To a lesser extent, they also produce words with shared vowel and manner in the coda, and words with the same onset. These results therefore provide some support for Kirtley et al.'s (1989) view that children represent words divided into their onsets and rimes. However, they also indicate that children sometimes spontaneously select words as sounding similar if they share a vowel and manner of articulation word-finally, which does not align with Kirtley et al.'s theory. These words should be treated as dissimilar since they do not share a rime.

Database analyses. It is important to examine the range of words available to children to complete this task, as it could be that children produce more rhyming words because more words rhyme with a given cue word than begin with the same initial consonant, for example. The four words used (bed, nut, shop and that) were all contained in a lexical database of child directed speech towards children under three years old (Carroll and Vousden, in preparation). The proportion of different types of neighbours was calculated for each word. Word sharing the initial consonant were relatively common (*Bed*: 42 words; *Nut*: 16 words; *Shop*: 19 words; *That*: 3 words). Body neighbours (in which the first consonant and vowel were shared) were relatively uncommon, and in each case less common than rime neighbours (*Bed*: 3 body neighbours, 6 rime neighbours; *Nut*: 1 body neighbour, 3 rime neighbours;

Shop: 2 body neighbours, 3 rime neighbours; *That*: 1 body neighbour, 8 rime neighbours).

Using these words, therefore, rime neighbours are less common than words sharing initial consonant (80 words sharing initial consonant, 20 words sharing rimes). However, in the child production data, rime neighbours were produced more often than words sharing initial consonant (110 words sharing initial consonant, 441 words sharing rimes). This deviation from the expected distribution is highly significant ($\chi^2(1, N = 651) = 1538.6, p < .001$).

In a further comparison of the database and child productions, rime neighbours are approximately 2.86 times as common as body neighbours in the database. In the production data, rime neighbours were produced over 15 times as often as body neighbours (441/28 = 15.75). Again, this deviation from the expected distribution is highly significant ($\chi^2(1, N = 496) = 25.5, p < .01$).

Out of all the CVC words that shared a vowel with the cue word but did not rhyme, just less than half (44.4%) had the same manner as the final consonant (*Bed*: 35 shared vowel words, 14 (40.0%) shared manner; *Nut*: 37 shared vowel words, 13 (35.0%) shared manner; *Shop*: 29 shared vowel words, 17 (58.6%) shared manner; *That*: 34 shared vowel words, 16 (47.0%) shared manner). In the production data, 79.1% of all words sharing vowels that did not rhyme shared manner word-finally. Again, this differs significantly from the expected distribution (($\chi^2(1, N = 269) = 67.0$, p < .01). Overall, the frequency of production of different word types is does not replicate what would be expected based on a child lexicon.

To summarise, children tended to produce rhyming words most commonly when asked to produce similar sounding words. They also frequently produced words that shared a vowel and manner word finally. This fits with the hypothesis that children use manner of articulation in word classification, especially word-finally, and with the hypothesis that vowels are more important in classification of word final consonants than word initial consonants. The results cannot be explained in terms of the frequency of different types of words in the child lexicon.

Forced Choice

It was anticipated that children would classify words sharing manner as sounding similar more often than unrelated words or words sharing place of articulation, and that this would be particularly true in the younger, pre-literate children. Matching on shared manner is likely to be particularly common word finally. Results could also be qualified by the neighbourhood density of the items used, as they were in Storkel's (2002) research.

It was anticipated that the particular words used and the order of presentation would not influence results. To assess this, two preliminary one way ANOVAs were carried out to assess whether there were significant main effects of Word Set (Storkel's words versus the new set) and Order of Presentation (body versus rime first). There were no significant effects for Word Set ($F(1,137) = 1.04$, $p = .309$, $\eta^2 = .01$) or for Order of Presentation ($F(1,137) = 1.21$, $p = .274$, $\eta^2 = .01$) and therefore these conditions are collapsed for further analyses.

A within subjects ANOVA was then carried out to examine whether Neighbourhood Density showed a significant main effect or interactions with the Relationship and Position variables. There was a main effect of Neighbourhood Density ($F(1,144) = 9.64$, $p = .002$, $\eta^2 = .06$), such that children were slightly more likely to say that words sounded similar if the target word was in a sparse phonological neighbourhood, but there were no significant interactions with this variable, and as neighbourhood density was not the focus of this study, this effect was not analysed further.

Figure 1 shows the pattern of 'sounds similar' responses across items. A mixed ANOVA was carried out to assess the effects of: featural relationship item ("Relationship") and position of the phonetic relationship ("Position"), together with age group ("Age"). Significant effects for Relationship ($F(4,576) = 384.98$, $p <.001$, $\eta^2 = .73$), and Position ($F(1,144) = 21.64$, $p <.001$, $\eta^2 = .13$) were found. The effect of Position occurred because children were more likely to classify words as sounding similar if the shared phonetic features were in the rime rather than the body of a word. Post-hoc contrasts were used to examine the effects of Relationship. There was a significant difference between each consonant relationship, with the exception of place of articulation and unrelated words.

There was a significant interaction between Relationship and Position ($F(4,368) = 23.82$, $p < .001$, $\eta^2 = .14$). This interaction is illustrated in Figure 1. In the body condition (Figure 1a), manner items tend to be classified as different, while in the rime condition (Figure 1b), manner items tend to be classified as similar. This impression was backed up by post hoc testing. Within the body condition, manner items did not differ significantly from place or unrelated items. Within the rime condition, manner items did differ significantly from place and unrelated items, which did not differ from one another.

Age group was not a significant predictor of overall performance ($F(2,144) = 0.17$, $p = .845$, $\eta^2 = .002$). As predicted, there was a significant interaction between Age Group and Relationship ($F(8,576) = 6.05$, $p < .001$, $\eta^2 = .08$). Older children were more likely to classify words sharing an initial or final phoneme as similar than the younger children were.

Item Analyses. The item analyses were carried out on data from the nursery and reception children only, since the Year 1 children had completed only Version A of the task. The results for the item analysis replicated the results for the subject analysis for the most part. There was no significant effect of old versus new stimuli ($F(1,78) = 1.18$, $p = .282$). The effect of neighbourhood density was significant ($F(1,20)$

$= 4.18$, $p = .045$, $\eta^2 = .07$). There was a significant effect of relationship ($F(4,60) = 212.11$, $p < .001$, $\eta^2 = .93$). Post-hoc tests showed that each type of relationship was significantly different from each other type of relationship with the exception of place and unrelated words. There was also a significant effect of position ($F(1,60) = 12.92$, $p =.001$, $\eta^2 = .18$) and a significant interaction between position and relationship ($F(4,60) = 7.22$, $p < .001$, $\eta^2 = .33$). This interaction was due to an increased proportion of 'sounds similar' responses to items which shared manner in the rime rather than the body. Within the rime items, there is no difference between the rate of responses for matching phonemes and for shared manner of articulation (Tukey's: $p = .13$), but both phoneme and manner differ significantly from place and unrelated ($p < .001$ in all cases). Within the body items the manner items are equivalent to the place and unrelated items (Tukey's: manner versus place $p = .85$; manner versus unrelated $p = .15$). No other interactions were significant.

There is a possible confound in examining manner of articulation with this sample of words. Nasal consonants only appear word-finally in the target words; in the rime condition. It is possible that the interaction between position and relationship is due to the fact that nasals occur only word-finally and nasals are a particularly similar-sounding set of consonants. In order to investigate this possibility, the analysis was repeated with the target words containing nasal consonants in the key position excluded. There remained a significant effect of relationship and position ($F(4,50) = 147.11$, $p < .001$, $\eta^2 = .92$ and $F(1,50) = 4.51$, $p = .04$, $\eta^2 = .08$), and a significant interaction between position and relationship ($F(4,50) = 2.69$, $p = .04$, $\eta^2 = .18$), though the strength of the relationship was reduced.

Overall, these results are in line with predictions. Word-initially, children are likely to classify words as similar if they share a full phoneme. Word-finally, words are classified as similar if they share a full phoneme or manner of articulation. In contrast to Storkel's (2002) results, the patterns were consistent across dense and sparse phonological neighbourhoods. While children were slightly more likely to say words sounded similar if they came from sparse neighbourhoods, this effect did not interact with position or relationship.

The results were also relatively consistent across age groups. The older children showed more consistent classification of the items sharing phonemes as similar, but they were no less likely to classify words sharing manner as similar. It may be that the older children are more consistently able to rate items sharing phonemes as most similar, but still recognise items sharing manner word-finally as similar. This partially substantiates Storkel's hypothesis that children will start relying more on phonemic similarity as they grow older.

Sound Families

The key difference between the Sound Families task and the Forced Choice task is that children are asked to match words to a group of cue words rather than an individual word. There were three further differences between this task and the Forced Choice task. First, only children from Nursery and Reception completed this task, which limits investigation of age related effects. Second, each participant completed either the Body or Rime condition. Finally, the presence of same and different vowels in this task also allowed an investigation of whether classification was different dependent on the vowel relationship. It was again predicted that children would be more likely to classify words as similar if they shared manner of articulation, and that this would be qualified by an interaction with position. There was a complication in that there were four types of shared vowel item and three types of different vowel item. Two sets of analyses were therefore carried out, one with the items sharing a vowel and one with items that contained a different vowel.

Shared Vowel Analysis. The results are illustrated in Figure 2. In a within subjects ANOVA comparing the four types of consonant relationships (same, manner, place and different) there is a significant effect of Relationship ($F(3,255) = 46.02$, $p < .001$, $\eta^2 = .35$) and a significant effect of Position ($F(1,85) = 10.02$, $p = .002$, $\eta^2 = .11$), together with a significant interaction between the two factors ($F(3, 255) = 5.52$, $p = .001$, $\eta^2 = .06$). Pairwise contrasts showed that there were significant differences between each of the conditions, with the exception of the place of articulation and the different conditions. As demonstrated in Figure 2, the pattern of responses differed between the body and rime conditions, with manner items being rated as similar in the rime condition but not the body condition. These results are in line with those of the Forced Choice task.

When including Age Group as a between subjects variable, there was no main effect of Age Group, but a significant interaction between Age Group and Relationship ($F(3,249) = 3.18$, $p= .025$, $\eta^2 = .04$). This interaction is shown in Figure 2. The older children differentiated between the phonetic relationships more consistently: they were more likely to accept words containing shared phonemes than any other words, while the younger children showed this preference to a lesser extent.

Different Vowel Analysis. The three types of items containing different vowels (manner different, place different and unrelated different) were then compared using a similar within subjects ANOVA. Mean scores suggested no strong relationships (body condition; manner: 0.34; place: 0.32; unrelated: 0.25. Rime condition; manner: 0.51; place: 0.37; unrelated: 0.32), and this was borne out by statistical analysis. There was no significant main effect of Relationship $F(2,166) = 1.93$, $p = .15$, $\eta^2 = .02$), though there was a significant effect of Position ($F(1,78) = 4.38$, $p =.04$, $\eta^2 = .05$), and the interaction between Position and Relationship only approached significance ($F(2,166) = 2.82$, $p = .06$, $\eta^2 = .03$). There is therefore no strong evidence for sensitivity to consonant features when two words contain different vowels.

Item Analysis. Similar results were found in an item analysis. When considering items that shared a vowel, there was a significant effect of Relationship ($F(3,28) = 39.87$, $p < .001$, $\eta^2 = .81$) and Position ($F(1,40) = 7.85$, $p = .009$, $\eta^2 = .22$) and a significant interaction between the two ($F(3,40) = 3.93$, $p =.018$, $\eta^2 = .30$). Post-hoc analyses showed that for the Body condition, the 'same' items were rated as similar more often than all of the other items, which did not differ from one another. For the Rime condition, the 'same' items and 'manner' items were both rated as similar and the 'place' and 'different' items were rated as different.

One of the item sets 'en' contained the family word 'jen' and the shared manner test word 'gem'. Since these two words are very similar, it could be that the effect of manner is carried by this item, so analyses were repeated removing this set of words. The results were similar: there was a significant effect of relationship ($F(3,24) = 34.46$, $p < .001$, $\eta^2 = .81$) and position ($F(1,24) = 6.59$, $p = .017$, $\eta^2 = .21$), though the interaction between the two was only marginally significant ($F(3,24) = 2.74$, $p =.065$, $\eta^2 = .26$). While post-hoc analyses could not be carried out with this item removed, proportion of 'sounds similar' ratings were similar for the 'same' items and the 'manner' items (same = .627, manner = .600), while the scores for the other items were lower (place = .215, different = .305), suggesting that the same pattern of results held once this item was removed.

When considering only different items there was no significant effect of relationship ($F(2,12) = 1.50$, $p = .297$) but a significant effect of position ($F(1,12) = 7.04$, $p = .038$, $\eta^2 = .54$), and the interaction between the two was not significant ($F(2,12) = 2.12$, $p = .201$).

There were stops and nasals in both the body and rime conditions. The particular manner or place of articulation of the sound family did not significantly affect the responses (Manner: $F(1,21) = 0.05$, $p = .82$, $\eta^2 <.01$; Place: $F(1,21) = 0.163$, $p = .69$, $\eta^2 = .01$).

The results are in line with the data from the forced choice task in that children are likely to classify words as sounding similar if they share manner of articulation word finally. Word initially, words tended to be classified as sounding similar if they shared an initial phoneme. Older children were more likely than younger children to classify words sharing a phoneme as similar. However, these effects were only present when words shared a vowel.

Individual Differences among the tasks

If the tasks used are assessing the consistency of underlying phonological representations, there should be significant associations between the tasks. Children who rate words sharing phoneme or manner as similar consistently should do so across the Forced Choice and Sound Families tasks, and potentially would also be more likely to produce more words in the Word Production task. In addition, it could be anticipated that vocabulary level and letter knowledge would be related to performance on these tasks, as Walley (1993) states that vocabulary growth and literacy instruction are major causes of lexical restructuring.

In order to examine consistency across the tasks, a number of summary variables were formed. In the classification tasks (Forced Choice and Sound Families), sensitivity measures were formed by adding all the 'sounds similar' responses given by each child and calculating two scores. The 'phoneme sensitivity' score represented the proportion of the 'sounds similar' responses that were in answer to items that shared a phoneme. The 'manner sensitivity' measure represented the proportion of 'sounds similar' responses that involved shared manner of articulation. Words that shared a full phoneme were included in this category, as words that share a full phoneme necessarily share manner of articulation within this phoneme. Therefore children using a purely manner based strategy would select items sharing phonemes as sounding similar in addition to items sharing purely manner of articulation. For the word production task, total number of words produced and total number of rhyming words produced were used as summary variables. These variables were then examined in a correlation matrix, shown in Table 3. Since 28 correlations are being examined, significance levels were designated at $p<.01$ rather than $p <.05$. Partial correlations were carried out to control for the possible confounding effects of age and number of terms in education. These correlations are parametric, since it is not possible to carry out non-parametric partial correlations.

If Walley's (1993) hypothesis is correct, one would predict that vocabulary would show a relatively specific association with the phonological processing measures. In fact, only the correlation with Forced Choice manner choices remained significant in the partial correlations. This suggests that the associations can be explained by broader factors such as age and educational experience.

On the other hand, letter knowledge shows a relatively specific association the phonological processing tasks, being significantly associated with the Forced Choice and Word Production tasks, even after age and education were controlled. This is in line with Walley's theory, but also in line with many theories that propose a close reciprocal link between letter knowledge and phonological processing (e.g. Carroll, Snowling, Hulme and Stevenson, 2003).

Within the three phonological processing tasks, the Sound Families Phoneme measure correlated well with both Forced Choice measures, as would be expected given the similar form of the two tasks.

There is some evidence that the Sound Families Manner measure does not tap the same skills as the Forced Choice measures, however. It shows no significant correlation with the Forced Choice task, and shows a significant negative correlation with Total Word Production, indicating that those children who were likely to produce many words in the Word Production task were the ones who were less likely to use manner based coding on the Sound Families task. Since production of many words is taken as an indicator of good phonological processing, this may indicate that manner based coding on the Sound Families task indicates relatively poor phonological coding.

Summary

The three experimental tasks presented here have demonstrated broadly similar results, despite both perception and production tasks being used. In each of the tasks, children showed a tendency to classify as similar words that shared manner word-finally. The fact that this occurs across tasks perhaps indicates that word-final manner based coding is a feature of phonological representations in children of this age. Each of the tasks also provided additional specific data. The Word Production task indicated that children tend to provide rhyming words when asked to produce words that sound like a cue word, and that they relatively rarely produce words starting with the same initial sound. The Forced Choice task indicated that the pattern of selecting as similar words that share manner word finally is present to a similar extent in three to six year old children, and there was little evidence that this tendency reduced with age. On the other hand, phonemic similarity was more consistently recognised by older children. The Sound Families task indicated that this effect was present only in words that shared a vowel. There was no consistent pattern of feature based classification in words that did not share a vowel.

On the Forced Choice and Sound Families tasks, older children were more likely to classify words sharing phonemes as sounding similar. However, older children still classified words sharing manner of articulation word-finally as similar. While phonemic classification became more common, manner classification did not become less common.

These tasks had two limitations. The first is that in each case the child was explicitly asked to consider which words sounded similar. Their responses may well be coloured by their view of what the experimenter expects them to say. For example, children may regard rhyming as the primary response to a question about sound similarities because of the prominence of rhymes and rhyming games in pre-school, rather than because they represent these two words in a similar way. A task which examines memory processes without explicit word comparisons would avoid this issue. The second limitation is the relatively small age range assessed. While Study 1 focuses on the age range in which phonological representations are said to change, it would be useful to have data from older children and adults, to clarify that young children do indeed respond in a qualitatively different way. Study 2 provides this data.

Study 2

Introduction

Much of the previous research in this area relies on standard phonological awareness tasks, which place demands on short term memory and speech perception, as well as the demands of reflecting on words out of context described above. It could be that the findings reflect task demands or perceptual processes rather than underlying representations. The tasks in Study 1 aimed to overcome these difficulties by using a range of different paradigms and providing as many prompts as necessary if children had memory difficulties.

Another way to avoid tasks in which children are asked to reflect upon speech is to assess confusion errors in a word learning task, in which children are not asked to make any comparisons between nonsense words. Treiman and Breaux (1982) carried out a study of this type. Children and adults were asked to learn sets of three nonsense names for animals. The children showed a tendency to confuse the names that were globally phonologically similar, while adults most commonly confused names that shared an initial phoneme.

The current study examines a modified version of this task in which three similarity dimensions are compared; shared manner of articulation, shared place of articulation and shared initial phoneme.

Method

Participants

The nursery, reception and year 1 children who took part in the Forced Choice task in Study One also took part in the Memory Confusions task in Study Two. In addition, a sample of 40 adults also took part. Most of these adults were psychology undergraduate students, who took part in the study to obtain course credit. The average age of participants was 20.5 years with a range of 17 years to 43 years. Twenty-eight females and 12 males took part in this study. Thirty-three participants spoke English as their first language with 5 participants having English as their second language and 2 participants being bilingual.

The Memory Confusion Task

This task uses a nonword learning paradigm to examine representations of new phonological forms (or names). It was based on the paradigm used by Treiman and Breaux (1982) in which children and adults are asked to learn the names of three animals and errors are recorded. While the basic paradigm remained the same, several aspects were altered. Within the Treiman and Breaux task the triplets contained pairs of words that shared an initial phoneme (such as *bis* and *boon*) and pairs (such as *deez* and *bis*) that were matched in terms of 'global phonological similarity' (based on acoustic similarity ratings). In the present study this task was modified to specifically examine the role of manner and place of articulation in memory. It was anticipated that participants would find pairs that they considered more similar more confusable in memory. Participants completed one of two alternative forms of the task, using the same toys but different sets of nonwords. Within each form, six triplets of nonwords were used. These triplets are shown in Appendix 3. One triplet contained a pair sharing a phoneme and one pair sharing manner, the second contained a pair sharing manner and a pair sharing place, and a final triplet contained a pair sharing a phoneme and a pair sharing place. This allowed a comparison between shared phonemes and consistent shared place or manner across both consonants within a nonword. Inevitably, given the relationships between the words, the final pair in each triplet shared some features. However, this pair did not share any phonemes or a consistent relationship in all consonants, and was therefore labelled the 'anomalous' pair. The vowel sounds within each triplet were all different. The order of presentation was counterbalanced.

To ensure that the certain animals were not visually or semantically more confusable than others, a pilot study was conducted in the manner suggested by Treiman and Breaux (1982), in which children were asked to say which animals went together. No consistent patterns emerged.

The nonsense words were presented as given names of toy animals, for example "the giraffe is called Mern". Participants heard the nonsense name and were asked to repeat it. Errors were corrected. The animals were then mixed up and the participants were given a name and had to point to the correct animal. Participants were given feedback after the three animals had been named, in the form: "you got this right, the giraffe *is* called Mern”. If the participant got all three names correct twice in a row that triplet was assumed to be learnt and they moved onto another triplet. The maximum number of trials for each set was six. This learning session was labelled Time A.

For the children, a delayed memory test (Time B) was carried out approximately twenty-four hours after the original learning session. The adults were tested only in a single session and so completed the retest approximately one hour after the first test. Participants were shown the animals and tested initially on whether they could remember the names without a reminder; if they could not, the procedure was repeated as above. The adults made very few errors on the Time B learning trials (only 5 errors in total across all participants) and so errors made on the initial retest were used instead of the learning trials for this group.

Error Analysis

Participants could make one of three types of response to each animal name. They could point to the correct animal, point to one of the other two animals or they could make some other kind of error, such as saying ‘I don’t know’ or not responding. Each answer was coded according to the type of response made. A confusion error (pointing to an incorrect animal) could be of four different types. Words sharing the same onset could be confused; words sharing manner of articulation or place of articulation could be confused; or words sharing a less consistent phonetic relationship (anomalous errors) could be confused. The total number of each type of confusion error over all of the word sets was calculated. If participants make errors at random, one would expect that 33.3% of their errors would be anomalous, and the remaining 66.7% errors would be distributed equally between manner confusions, phoneme confusions and place confusions (e.g. 22.2% of the errors would fall into each of these categories).

Results and Discussion

This task uses a learning task to examine memory processes. If phonetic features form part of representations, words sharing these features would be more likely to be confused. Confusions formed 63% of all of the errors made on the memory task. The key question of interest is whether children and adults make one type of phonological confusion more than another type. The numbers of each type of confusion across word sets at time A and time B for each age group is shown in Table 4.

Child Data

Each age group, except the reception group, shows an overall pattern of responses which differs significantly from chance at Time A (Nursery: $\chi^2(3, N = 495) = 24.32$, $p < .01$; Reception: $\chi^2(3, N = 386) = 2.37$, $p = .50$; Year 1: $\chi^2(3, N = 367) = 21.85$, $p < .01$) and all groups differ significantly from chance at Time B (Nursery: χ^2(3, N =

353) = 44.24, $p < .01$; Reception: $\chi^2(3, N = 231) = 27.25$, $p <.01$; Year 1: $\chi^2(3, N = 358) = 14.99$, $p < .01$).

Overall, there is a trend for the proportion of phoneme based confusions to increase with age and the proportion of manner based confusions to decrease with age in childhood, both at Time A and at Time B.

At Time A, the patterns of responses were as follows. The Nursery children made more manner errors ($\chi^2(1, N = 495) = 15.16$, $p <.01$) than expected by chance, but expected levels of phoneme errors ($\chi^2(1, N = 495) = 0.05$, $p = .83$). The Reception children showed expected levels of manner errors ($\chi^2(1, N = 386) = 1.28$, $p = .26$) and phoneme errors ($\chi^2(1, N = 386) = 0.21$, $p = .64$). The Year 1 children showed expected levels of manner errors ($\chi^2(1, N = 367) = 0.49$, $p = .49$), but more phoneme errors ($\chi^2(1, N = 367) = 14.62$, $p <.01$). The children therefore show a pattern of moving from manner based errors towards phoneme based errors with increasing age.

Similar findings occurred at Time B. The Nursery children made more manner errors ($\chi^2(1, N = 353) = 26.97$, $p <.001$) than expected by chance, but expected levels of phoneme errors ($\chi^2(1, N = 353) = 1.83$, $p = .18$). The Reception children showed higher than expected levels of manner errors ($\chi^2(1, N = 231) = 6.96$, $p = .01$) and expected levels of phoneme errors ($\chi^2(1, N = 231) = 2.34$, $p = .13$). The Year 1 children expected levels of manner errors ($\chi^2(1, N = 358) = 0.32$, $p = .57$), but more phoneme errors ($\chi^2(1, N = 358) = 9.66$, $p <.001$).

These data are in line with predictions that pre-literate children would tend to use feature-based coding when learning new words, while older children would tend to use phoneme based coding. The middle age group do not show either pattern clearly, perhaps indicating that they are in transition.

Adult Data

The overall distribution of responses differed significantly from chance at Time A ($\chi^2(3, N = 97) = 59.99$, $p < .01$) and at Time B ($\chi^2(3, N = 71) = 21.08$, $p < .01$).

At Time A, the adults seem to make similar levels of manner and phoneme based errors, and they make very few place based or anomalous errors. At Time B, in contrast, phoneme based errors are most common. This analysis is backed up statistically. At Time A, the adults showed both more manner errors and more phoneme errors than expected by chance (Manner: $\chi^2(1, N = 97) = 22.56$, $p <.001$; Phoneme: $\chi^2(1, N = 97) = 20.30$, $p <.001$). At Time B, the adults showed expected levels of manner errors ($\chi^2(1, N = 71) = 0.40$, $p = .53$), but higher than expected levels of phoneme based errors ($\chi^2(1, N = 71) = 16.49$, $p <.001$). The adults therefore seem to show a pattern of using manner based coding when first dealing with new words, but using predominantly phoneme based coding in longer term recall.

The adults showed different patterns at the two test points; at time A, both manner and phoneme based confusions were common, while at the retest, the majority of confusions were phoneme based. A wide range of studies indicate that adults accord a special status to word initial sounds in word recognition (e.g. Nooteboom, 1981), which would be in line with the high number of initial phoneme confusions at Time B. At Time A, it is possible that because words are being held in short-term memory, confusions occur due to both acoustic and phonemic similarities between the words. At Time B, the adults had not heard the words for around an hour and were relying on long-term memory to a greater extent.

The children did not show this discrepancy between the two test points. This may be because they store newly learnt words differently from adults (possibly using acoustic information to a greater extent and phonemic information to a lesser extent). Alternatively, it could be because the children had not learnt the words as effectively as the adults at Time B and continued to rely on short-term memory during the learning trials. In other words, it is difficult to disentangle, on the basis of this data, whether children and adults form fundamentally different underlying representations for newly learnt words, or whether differences in cognitive processing cause different patterns of memory errors.

General Discussion

These studies aimed to examine the use of phonetic features in the classification and representation of words by children and adults. In particular, we were interested in whether use of phonetic features in tasks varied according to the age of the child, the nature of the task or the position of the feature in question. Four tasks were used, with all giving some evidence for sensitivity to shared manner of articulation. This was particularly common word finally. In addition, all tasks showed sensitivity to shared phonemes in both initial and final position. There was evidence of increasing use of phonemic information with increased age, though sensitivity to shared manner was shown throughout the age range.

The results provide evidence that children are sensitive to sound similarities beyond the level of the phoneme, and that in particular they are sensitive to manner similarities word-finally. This provides a possible explanation of Byrne and Fielding-Barnsley's (1993) finding that when preschool children are asked to decide whether two words share sounds, they often regard words that do not share any phonemes (but share phonetic features) as sounding similar.

With increasing age, children tended to use phonemic information more consistently. In the older age groups, items sharing phonemes were rated as similar more often and phoneme based errors were more common in the memory task. However, this was not consistently accompanied by a reduction in rating words sharing manner as sounding similar. Even adults showed high numbers of manner based errors on first phase of the nonword learning task. Since items sharing manner of articulation are acoustically similar, this is perhaps unsurprising. Rather than a drop in sensitivity to shared manner over time, these results seem to indicate that children begin to become more flexible in the phonological information they use with increasing age, perhaps as they come to represent sounds more consistently across words.

There were associations between letter knowledge and performance on the tasks that remained after controlling for age and educational level, indicating a possible link between performance on these tasks and literacy development. This is in line with Walley's (1993) theory that literacy tuition alters phonological representations, though her idea that vocabulary would be another important factor was not substantiated.

No task in these studies indicates sensitivity to shared place of articulation. There are at least two possible reasons for this. First, the examples used here employed a broad categorisation of place of articulation (anterior, coronal or dorsal) rather than more specific categories. It is possible that this masked sensitivity to place of articulation. Alternatively, it could be that children are sensitive to place of articulation only in phonemes that share manner. Previous studies that have shown sensitivity to place of articulation (e.g. Snowling et al., 1994) have held manner of articulation constant.

The results provide support for the idea that word-final consonants are represented in less detail than word-initial consonants, as outlined in the introduction. Less fully represented words would be matched less reliably, and additional phonological information could help to increase the reliability of classification. This could take the form of shared vowels, as in the Forced Choice and Sound Families tasks, or shared features both word-initially and word-finally, as in the Memory Confusions task.

These studies extend previous research by showing that young children are sensitive to both word structure and manner of articulation. They also show the importance of task demands in the interpretation of results: data from Study 1 shows only an increase in phoneme based classification with age, while the data from Study 2 shows both an increase in phoneme classification and a decrease in manner based classification. The results lead one to the conclusion that children do change in the way that they process words with across this age range, and the consistency across tasks suggests this may be linked with changes in phonological representations themselves.

Acknowledgements

This research was carried out with the help of a grant to the first author from the Economic and Social Research Council, number RES-000-22-1209. We thank Prof. Maggie Snowling, Prof. Gordon Brown and James Clyne for their comments on earlier drafts of this manuscript, and we thank the schools, children and parents who helped with the data collection.

References

Bailey, T. M., & Hahn, U. (2005). Phoneme similarity and confusability. *Journal of Memory and Language, 52*, 339-362.

Bailey, T. M., & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development, 17*, 1265-1282.

Ballem, K. D., & Plunkett, K. (2005). Phonological specificity in children at 1;2. *Journal of Child Language, 32*, 159-173.

Byrne, B., & Fielding-Barnsley, R. (1993). Recognition of phoneme invariance by beginning readers: Confounding effects of global similarity. *Reading and Writing, 5*, 315-324.

Cardoso-Martins, C. (1994). Rhyme perception: global or analytic? *Journal of Experimental Child Psychology, 57*, 26-41.

Carroll, J. M., & Snowling, M. J. (2001). The effects of global similarity between stimuli on performance on rime and alliteration tasks. *Applied Psycholinguistics, 22*, 327-342.

Carroll, J. M., Snowling, M. J., Hulme, C., & Stevenson, J. (2003). The development of phonological awareness in pre-school children. *Developmental Psychology, 39*, 913-923.

Carroll, J. M., & Vousden, J. I. (in preparation). A lexicon of child directed speech.

Creel, S. C., & Dahan, D. (2010). The effect of the temporal structure of spoken words on paired-associate learning. *Journal of Experimental Psychology: Learning, Memory and Cognition, 36*, 110-122.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in Infants. *Science, 171*, 303-306.

Fennell, C. T. & Werker, J. F. (2003). Early word learners' ability to access phonetic detail. *Language and Speech, 46*, 245-264.

Ferguson, C. A., & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language, 51*, 419-439.

Gerken, L. A., Murphy, W. D., & Aslin, R. N. (1995). Three- and four-year-olds' perceptual confusions for spoken words. *Perception and Psychophysics, 57*, 475-486.

Grunwell, P. (1997). Natural Phonology. In M. J. Ball & R. D. Kent (Eds.), *The New Phonologies*. San Diego: Singular Publishing Group.

Jusczyk, P., Goodman, M. B., & Bauman, A. (1999). Nine-months-olds' attention to sound similarities in syllables. *Journal of Memory and Language, 40*, 62-82.

Kirtley, C., Bryant, P. E., MacLean, M., & Bradley, L. (1989). Rhyme, rime, and the onset of reading. *Journal of Experimental Child Psychology, 48*, 224-245.

Locke, J. L. (1983). *Phonological Acquisition and Change*. London, UK: Academic Press.

MacNeilage, P. F. & Davis, B. L. (2000). On the origin of internal structure of word forms. Science, 288, 527-531.

Magnuson, J.S., Tanenhaus, M. K., Aslin, R. N., & Dahan, D. (2003) The time course of spoken word recognition: Studies with artificial lexicons. *Journal of Experimental Psychology: General, 132*, 202-227.

Mani, N. & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language, 57*, 252-272.

Morais, J. (2003). Levels of phonological representation in skilled reading and in learning to read. *Reading and Writing, 16*, 123-151.

Morton, J., & Frith, U. (1993). What lesson for dyslexia from Down's Syndrome? Comments on Cossu, Rossini and Marshall (1993). *Cognition, 48,* 289-296.

Nooteboom. (1981). Lexical retrieval from fragments of spoken words: Beginnings vs endings. *Journal of Phonetics, 9*, 407-424.

Piske, T. (1997). Phonological organization in early speech production: Evidence for the importance of articulatory patterns. *Speech Communication. Special Issue: Speech Production: Models and Data, 22*, 279-295

Shockey, L., & Bond, Z. S. (1980). Phonological processes in speech addressed to children. *Phonetica, 37*, 267-274.

Singh, S., & Woods, D. R. (1971). Perceptual structure of 12 American English Vowels. *Journal of the Acoustical Society of America, 49*(6), 1861-1866.

Singh, S., Woods, D. R., & Becker, G. M. (1972). Perceptual structure of 22 English vocalic consonants. *Journal of the Acoustical Society of America, 52*(6), 1698-1713.

Snowling, M. J., Hulme, C., Smith, M. M., & Thomas, G. (1994). The effects of phonetic similarity and list length on children's sound categorisation performance. *Journal of Experimental Child Psychology, 58*, 160-180.

Sosa, A. V. & Stoel-Gammon, C. (2006). Patterns of intra-word phonological variability during the second year of life. *Journal of Child Language, 33*, 31-50.

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388*, 381-382.

Storkel, H. L. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language and Hearing Research, 44*, 1321-1337.

Storkel, H. L. (2002). Restructuring of similarity neighbourhoods in the developing mental lexicon. *Journal of Child Language, 29*, 251-274.

Swingley, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science, 8*, 432-443.

Swingley, D. (2009). Onsets and codas in 1.5 year olds' word recognition. *Journal of Memory and Language, 60*, 252-269.

Treiman, R., & Breaux, A. M. (1982). Common phoneme and overall similarity relations among spoken syllables: their use by children and adults. *Journal of Psycholinguistic Research, 11*(6), 569-598.

Vihman, M. M., Nakai, S., DePaolis, R. A., & Halle, P. A. (2004). The role of accentual pattern in early lexical representation. *Journal of Memory and Language, 50*, 336-353.

Walley, A. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review, 13*, 286-350.

Wechsler, D., Rust, J., & Golombok, S. (2003). *Wechsler Preschool and Primary Scale of Intelligence*. London, UK: Psychological Corporation.

White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language, 59*, 114-132.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131*(1), 3-29.

Footnotes

[1] An additional task, phonological priming, was presented to the children, but it is not included in this study because it showed a lack of reliability with this sample.

[2] No measure of whether this task transferred between the visual and auditory modalities was taken. It was treated as an informal introduction.

[3] Analyses are presented by items, for ease, but similar results are found when analysing the word by child.

Appendix 1

a) Practice items for the Forced Choice task

Sounds similar	Sounds different
Fish	Neck
Wish	Tag
Vich (used twice)	Rug
Biff	Dip
	Dime

b) Test items for the Forced Choice task

Density	Target word	Body condition			Rime condition			Unrelated
		Phoneme	Manner	Place	Phoneme	Manner	Place	
Set A: Original Words Storkel (2002)								
Sparse	**tug**	tough	bus	sun	hug	mud	young	mum
	dime	died	bite	night	mime	mine	wipe	mice
Dense	**tap**	tan	pan	sang	map	fat	ham	fan
	tin	tick	pig	sick	win	him	hit	fig
Set B: New Words								
Sparse	**dog**	doll	bomb	shone	fog	job	wrong	moth
	pen	peg	deck	web	hen	them	said	neck
Dense	**but**	bung	gum	fudge	nut	rug	run	sum
	phone	folk	soap	poke	tone	comb	coat	loaf

Appendix 2 Target families and test items for the Sound Families task

Body Condition	Practice session	Session 1		Session 2 ..	
Target group	**ri..**	**bea..**	**ti..**	**ma..**	**do..**
	rich	bead	tick	made	dock
	rick	beak	tiff	mage	dog
	rid	bean	tin	mail	doll
	rip	beam	tip	mate	dots
	ring	beach	tizz	maze	doss
Control items	rig	beep	tig	make	dom
	rim	beat	tim	main	don
	ring	beef	till	maim	dov
	ridge				
Manner and same vowel		team		name	
Place and same vowel			sit		nod
Different and same vowel			whip	cake	
Manner and different vowel			card		peel
Place and different vowel		food		part	
Different and different vowel	gus	shirt			fool
	pot				
	ham				
	maze				

a) Body condition

Rime Condition	Practice Items	Session 1		Session 2	
Target group	**.. us**	**..at**	**..ot**	..**an**	..**en**
	bus	bat	cot	can	den
	fuss	cat	dot	dan	hen
	jus	chat	got	fan	jen
	pus	sat	pot	than	ten
	gus	that	what	van	when
Control items	us	fat	hot	man	ken
	gus	mat	not	pan	pen
	mus	pat	lot	ran	men
	thus				
Same vowel and manner		map			gem
Same vowel and place		gas		cash	
Same vowel and different			cough	wrap	
Different vowel and manner			tap	sing	
Different vowel and place			guess		hit
Different vowel and different	swap	ring			wake
	lamb				
	hop				
	rock				

b) Rime condition

Appendix 3. Nonsense words triplets used in memory confusions task.

Relationship	Triplet 1a:	Triplet1b:	Triplet 1c:
	Pim (pɪm)	Mern (mɜn)	Boop (bup)
	Tane (teɪn)	Moab (məʊb)	Dut (dʌt)
	Pooss (pus)	Vit (vɪt)	Mave (meɪv)
Shared manner	pim tane		boop dut
Shared phoneme	pim pooss	mern moab	
Shared place		mern vit	boop mave
Anomalous pair	tane pooss	moab vit	mave dut
Relationship	Triplet 2a:	Triplet 2b:	Triplet 2c:
	Fep (fɛp)	Verg (vɜg)	Theen (θin)
	Thak (θæk)	Vol (vɒl)	Soam (səʊm)
	Thoaz (θəʊz)	Beng (bɛŋ)	Doob (dub)
Shared manner	fep thak		theen soam
Shared phoneme	thak thoaz	verg vol	
Shared place		verg beng	doob soam

Table 1: The experimental tasks used in the paper

Task Name	Activity	Example item	Participants
Word Production	Children produce a word that sounds like a given word.	"Can you think of any words that sound like *shop*?"	Nursery Reception
Forced Choice	Children decide whether words sound similar to a target word.	"Does *bus* sound like *tug*?"	Nursery Reception Year 1
Sound Families	Children decide whether a word belongs within a 'family' of words.	"Here are some words that sound a bit the same: *bean, beam, beach, beat, beef*. Does *team* belong in the family?"	Nursery Reception
Memory Confusions (Study 2)	Participants learn nonsense names for animals.	"Which animal is called *Mave*?"	Nursery Reception Year 1 Adults

Table 2: Rates of occurrence of phonologically related CVC words in the production task

Coda properties						
	Same	Differs in voicing	Shared manner	Shared place	No shared features	Total
Shared vowel	441 (76.7)	13 (2.3)	106 (18.4)	11 (1.9)	4 (0.7)	575
Different vowel	52 (38.8)	8 (6.0)	39 (29.1)	26 (19.4)	9 (6.7)	134
Onset properties						
	Same	Differs in voicing	Shared manner	Shared place	No shared features	Total
Shared vowel	28 (4.9)	44 (7.6)	90 (15.7)	129 (22.4)	284 (49.4)	575
Different vowel	82 (61.2)	8 (6.0)	11 (8.2)	8 (6.0)	25 (18.6)	134

Note: percentages of words produced are in parentheses.

Table 3: Correlations between performance on the tasks

	Vocabulary	Letters	SF phoneme	SF manner	FC phoneme	FC manner	Total production
Age							
Vocabulary	-						
Letters	.501**	-					
SF phoneme	.039 (.74)	.173 (.13)	-				
SF manner	-.145 (.20)	.077 (.50)	.578**	-			
FC phoneme	.173 (.12)	.240 (.03)	.466**	.146 (.23)	-		
FC manner	.233 (.03)	.342 **	.483**	.120 (.33)	.899**	-	
Total Prod.	.227 (.01)	.295 **	-.045 (.69)	-.328 **	-.044 (.69)	.023 (.83)	-
Rhyme Prod	.128 (.17)	.276 **	-.082 (.47)	-.237 (.03)	-.006 (.96)	.031 (.78)	.546**

Note: raw correlations are shown above the diagonal. Correlations controlling for age and education are shown below the diagonal.

Correlations that are significant at $p<.01$ are indicated by a double asterisk (**). Other probability levels are provided in parentheses.

Table 4: Percentage occurrence of the different types of memory confusions.

	Shared manner	Shared phoneme	Shared place	Anomalous	Total
Time A					
Nursery (n = 55)	29.49 (146)	21.82 (108)	14.74 (73)	33.40 (169)	100 (495)
Reception (n=53)	24.61 (95)	21.24 (82)	23.58 (91)	30.57 (118)	100 (386)
Year 1 (n = 58)	20.71 (76)	30.52 (112)	14.44 (53)	34.33 (126)	100 (367)
Adults (n = 40)	42.27 (41)	41.24 (40)	6.19 (6)	10.31 (10)	100 (97)
Time B					
Nursery (n = 55)	33.71 (119)	25.21 (89)	22.10 (78)	18.98 (67)	100 (353)
Reception (n=53)	29.44 (68)	26.41 (61)	26.84 (62)	17.32 (40)	100 (231)
Year 1 (n = 58)	23.46 (84)	29.05 (104)	22.34 (80)	25.14 (90)	100 (358)
Adults (n = 40)	25.35 (18)	42.25 (30)	8.45 (6)	23.94 (17)	100 (71)

Note: raw number of errors is shown in parentheses

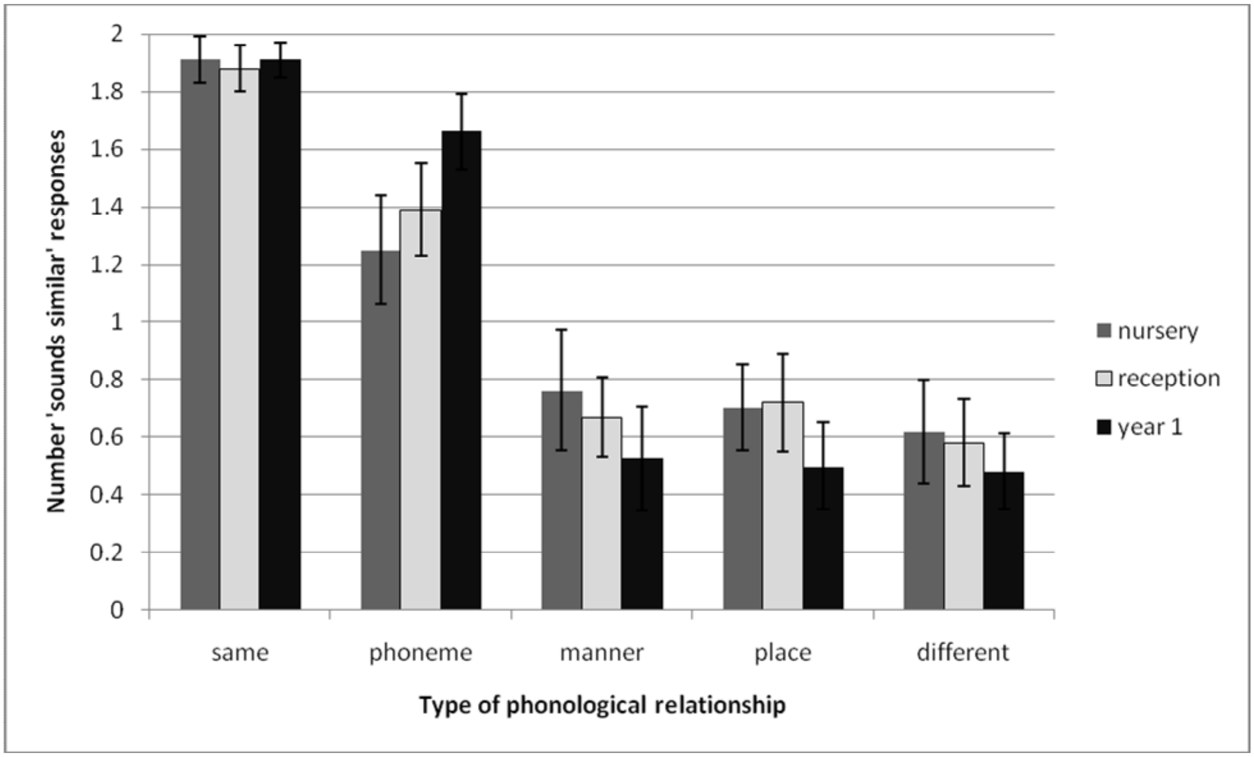
Figure Captions

Figure 1: Effects of age, type and position of sound similarity in the Forced Choice task

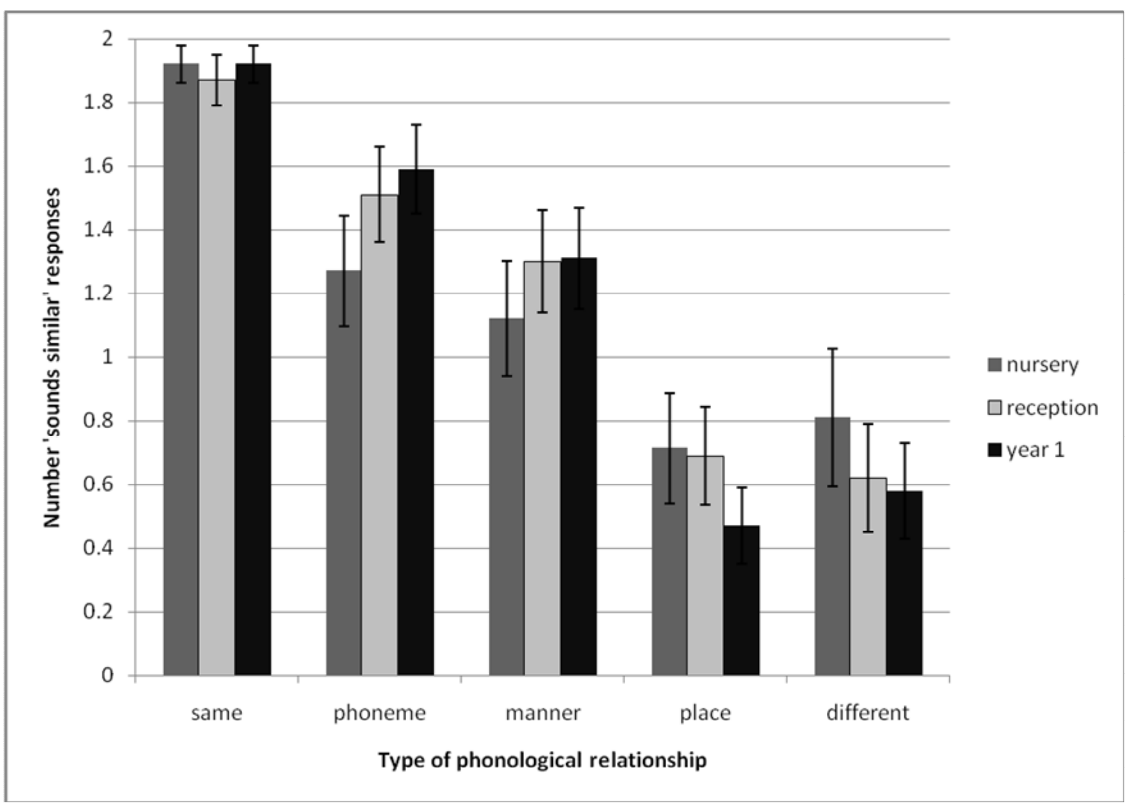
Note: the error bars show 95% confidence intervals for each data point

Figure 2. Effects of age, type and position of sound similarity in the Sound Families task

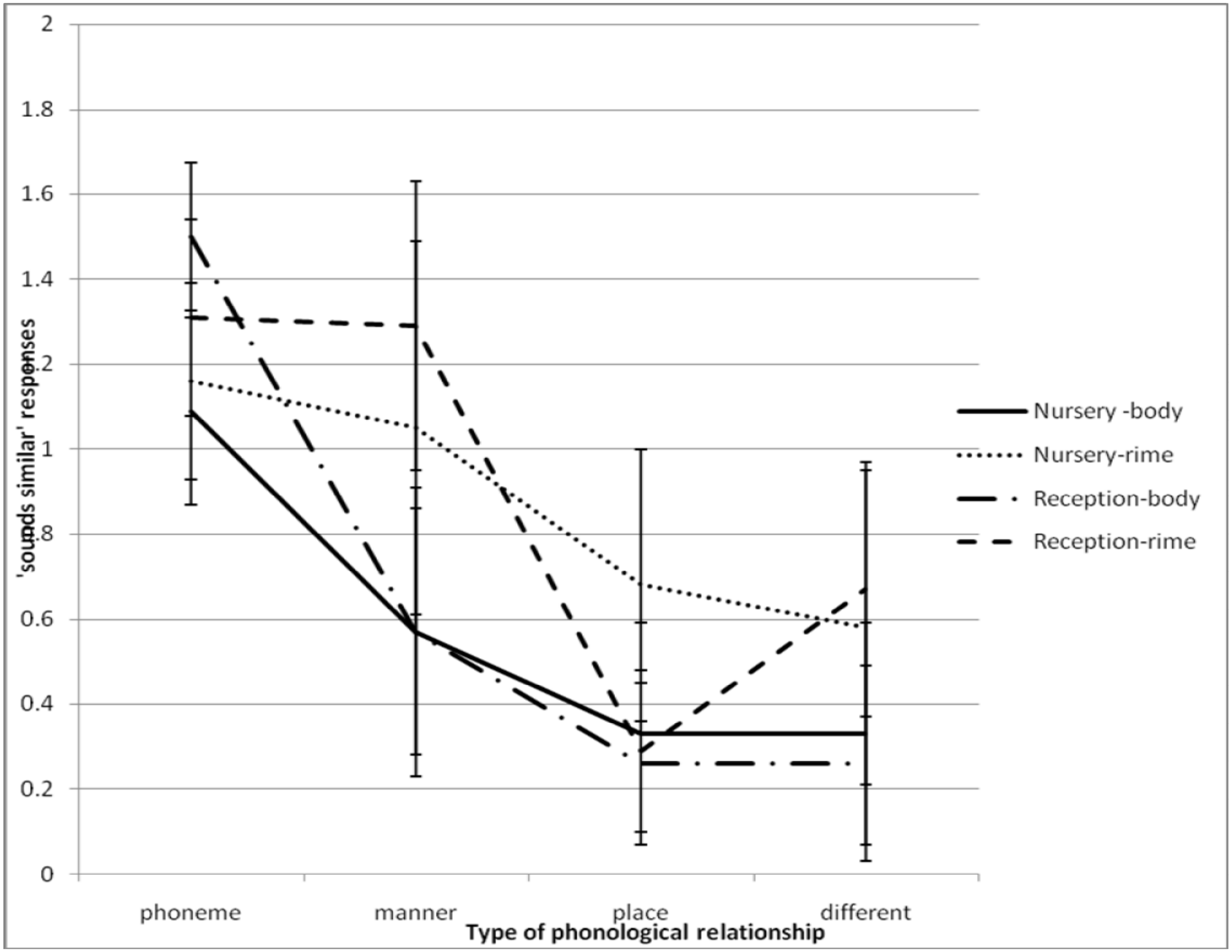
Note: the error bars show 95% confidence intervals for each data point



a) Responses in the body condition



b) Responses in the rime condition



2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
'sounds similar' responses
phoneme
manner
place
different
Type of phonological relationship
Nursery -body
Nursery-rime
Reception-body
Reception-rime