

**DEVELOPMENT OF PHONOLOGICAL
REPRESENTATIONS IN YOUNG CHILDREN**

A thesis submitted to the University of Manchester for the
degree of Doctor of Philosophy in the Faculty of Medical and Human
Sciences

2015

Stephanie Ainsworth

School of Psychological Sciences

Table of Contents

List of Tables.....	8
List of Figures	10
Abstract	12
Declaration	13
Copyright Statement	13
Acknowledgements	14
The Author	14
CHAPTER 1: INTRODUCTION	15
Thesis overview	16
Accessibility versus emergence: Are children’s phonological representations adult-like from infancy?.....	17
Precipitation of the phoneme: Does phonemic representation emerge naturally in the absence of literacy?	22
Similarity judgements as a probe into the lexicon.....	26
A connectionist approach to the development of phonological representations	29
Connectionist models of reading	29
Connectionist models of spoken language	31
Research aims.....	33
Acknowledgement of contribution of other authors.....	35
CHAPTER 2: SEGMENTAL PHONOLOGY HAS ITS ORIGINS EXPLICITLY IN LETTERS BUT IMPLICITLY IN WORDS	36
Abstract.....	37
Introduction	37
Key accounts of PR development.....	38

The Accessibility Account.....	38
The Emergent Account	39
i) Vocabulary growth drives PR segmentation which is a prerequisite for PA	40
ii) Children’s phonological representations become phonemic only once they have learnt the mappings between graphemes and phonemes	40
iii) Lexical restructuring occurs in the absence of literacy but letter knowledge is needed for explicit awareness of phonemes	41
Predictions made by the key theories of phonological development	41
Similarity classification as an implicit measure of phonological representation	44
Method.....	45
Participants	45
Procedure	45
Materials	46
Implicit measures of phonological representation	46
Similarity based classification tasks	46
Global similarity matching	47
Mispronunciation reconstruction.....	51
Pseudoword similarity.....	51
Mispronunciation conflict	52
Incomplete word.....	52
Implicit rhyme.....	52
Explicit measures of phonological representation	53
Explicit rhyme.....	53
Blending.....	53
Phoneme isolation	53
Background language measures	53
Letter-sound knowledge.....	53
Vocabulary	53
Counterbalancing	54

Results	55
Data screening	55
Summary of children's performance	55
How are PR segmentedness, phonological awareness, vocabulary and letter-sound knowledge related to one another?	56
Do letter-sound mappings need to be learnt before phoneme level phonological representations emerge?	61
Do we need letter-sound knowledge for explicit phonological awareness?	62
Discussion	64
The role of vocabulary growth	64
The role of letter-sound knowledge	65
Conclusion	67
CHAPTER 3: PHONEMES EMERGE FROM WORDS BUT ARE NOTICED THROUGH LETTERS	68
Abstract	69
Introduction	70
Theoretical accounts of phonological development	70
Predictions made by key theories of phonological development	73
Method	77
Participants	77
Procedure	77
Materials	77
Implicit measures of phonological representation	78
Similarity based classification tasks	78
Global similarity matching	79
Mispronunciation reconstruction	79
Pseudoword similarity	79
Mispronunciation conflict	79
Incomplete word	80

Implicit rhyme	80
Explicit measures of phonological representation	80
Explicit rhyme	80
Blending.....	81
Phoneme isolation.....	81
Background language measures	81
Letter-sound knowledge	81
Vocabulary.....	81
Counterbalancing.....	81
Results	83
Data screening	83
Developmental performance on implicit and explicit measures of phonological representation	83
Performance on measures of rime versus phoneme level representation.....	85
The relative role of vocabulary growth and letter-sound knowledge in the development	
of PR segmentedness and explicit phonological awareness.....	90
Discussion.....	95
The development of implicit and explicit knowledge of phonological representation	95
The role of vocabulary growth and letter-sound knowledge	96
Conclusion.....	98
CHAPTER 4: GLOBAL VERSUS PHONEMIC SIMILARITY: EVIDENCE IN SUPPORT OF MULTI-LEVEL REPRESENTATION	99
Abstract.....	100
Introduction	101
Global versus segmental representation	101
Developmental changes in global sensitivity	103
Method.....	106
Participants	106
Procedure	106

Materials	106
The mispronunciation reconstruction task.....	107
Global similarity matching	107
Results	109
How do response patterns change over development?	109
Sensitivity to phonemic versus global similarity.....	111
How does sensitivity to phonemic and global similarity change over development?	112
Discussion.....	117
Conclusion	120
CHAPTER 5: EVIDENCE OF EMERGING SEGMENTAL STRUCTURE WITHIN A NEURAL NETWORK MODEL TRAINED USING REAL SPEECH DATA	121
Abstract.....	122
Introduction	123
Computational models of spoken language.....	125
Method.....	127
Model overview	127
Representations.....	128
Temporal dynamics	129
Training.....	130
Results	132
How accurate is the model?	132
How well does the model generalize its knowledge to unseen words?	133
Is there evidence of emerging segmentation within the model’s hidden representations?	134
Similarity metric	135
Sensitivity to the number of shared phonemes and global similarity.....	138
Discussion.....	141
Does phonemic representation emerge within a preliterate model?	141
Conclusion	146
CHAPTER 6: GENERAL DISCUSSION.....	147

Summary of thesis findings	148
Theoretical implications	149
Accessibility versus emergence: Are children’s phonological representations adult-like from infancy?.....	149
Precipitation of the phoneme: Does phonemic representation emerge naturally in the absence of literacy?.....	152
Sensitivity to global versus phonemic similarity.....	154
Methodological implications	156
Similarity judgements as a probe into the lexicon.....	156
Connectionist models of spoken language using real speech data	158
Implications for educational practice	159
Conclusions and future directions	161
REFERENCES	163
APPENDICES	175
Appendix 1: Items for the mispronunciation conflict task	175
Appendix 2: Items for the mispronunciation reconstruction task	176
Appendix 3: Items for the pseudoword similarity task	177
Appendix 4: Items for the incomplete word task	178
Appendix 5: Items for the implicit and explicit rhyme tasks	179
Appendix 6: Items for the blending and phoneme isolation tasks	180
Appendix 7: Items for the letter-sound knowledge task	181
Appendix 8: Matching statistics for tasks where the two and one phoneme choices were matched for global similarity	182
Appendix 9: Matching statistics for frequency	183
Appendix 10: Matching statistics for phonotactic probability: positional segment average	184
Appendix 11: Matching statistics for phonotactic probability: biphone average	185
Appendix 12: Words used to train and test the neural network model	186

Word Count: 45,031

List of Tables

Table 2.1. Summary of predictions made by the four key theories of phonological development	42
Table 2.2. Summary of measures	48
Table 2.3. Loadings for the implicit PR, explicit PR and vocabulary PCA components	58
Table 2.4. Partial correlations between implicit PR, explicit PR, vocabulary and letter-sound knowledge when controlling for age	60
Table 2.5. Regression model with implicit PR as the outcome variable and age, explicit PR, vocabulary and letter-sound knowledge as the predictors	60
Table 2.6. Regression model with explicit PR as the outcome and age, implicit PR, vocabulary and letter-sound knowledge as the predictors	61
Table 2.7. Percentage of children within the low letter knowledge group who succeeded on each measure of implicit PR and explicit PR.....	63
Table 3.1. A summary of the relative roles of vocabulary growth and letter-sound knowledge on phonological representation and phonological awareness at the rime and phoneme level	75
Table 3.2. Number of tasks completed at each time point by the four children with incomplete data profiles	83
Table 3.3. Task loadings for the six principal components.....	88
Table 3.4. Partial correlations between the six key variables while controlling for age.....	90
Table 3.5. GEE results for model 1 with implicit PR rime as the outcome variable and explicit PR rime, vocabulary, letter-sound knowledge and age as predictors	92
Table 3.6. GEE results for model 2 with implicit PR phoneme as the outcome and explicit PR phoneme, vocabulary, letter-sound knowledge and age as predictors	92
Table 3.7. GEE results for model 3 with explicit PR rime as the outcome variable and implicit PR rime, vocabulary, letter-sound knowledge and age as predictors	93
Table 3.8. GEE results for model 4 with explicit PR phoneme as the outcome variable and implicit PR phoneme, vocabulary, letter-sound knowledge and age as predictors.....	93
Table 4.1. Phonemic and global sensitivity change scores assigned to longitudinal response patterns	113

Table 5.1. Predictions made by key theoretical accounts of phonological development about the emergence of phonemic representation	124
Table 5.2. Comparison of mean agreement between repeated words within the training set and mean percentage accuracy for all words at different developmental points	133
Table 5.3. Mean percentage accuracy for 94 untrained test words at different developmental points	134

List of Figures

Figure 2.1. Implicit measures of phonological representation.....	49
Figure 2.2. Performance on implicit PR measures with a chance level of a) 25% and b) 50%	57
Figure 2.3. Performance on explicit measures of phonological representation	58
Figure 2.4. Clustered bar chart showing increasing performance on all 4 measures with age.	60
Figure 3.1. Causal pathways predicted by four theoretical accounts of phonological development	74
Figure 3.2. Performance on implicit measures of phonological representation – mispronunciation reconstruction, incomplete word, implicit rhyme, mispronunciation conflict and pseudoword similarity	84
Figure 3.3. Performance on explicit measures of phonological representation	85
Figure 3.4. Performance on implicit PR measures – incomplete word, mispronunciation reconstruction, mispronunciation conflict and pseudoword similarity with a chance level of a) 25% and b) 50%	86
Figure 3.5. Performance on explicit PR measures split into rime level versus phoneme level items	87
Figure 3.6. Developmental progression across the six variables of theoretical interest	89
Figure 3.7. Summary of the GEE results	94
Figure 4.1 Children’s response profiles across four different time points with adult performance included for comparison	110
Figure 4.2. Children’s response profiles across four cross-sectional groups with adult performance included for group comparison	110
Figure 4.3. Developmental plot of children’s sensitivity to phonemes and global similarity from the longitudinal data	114
Figure 4.4. Developmental plot of children’s sensitivity to phonemes and global similarity from the cross-sectional data.....	115
Figure 5.1. Architecture of the feedforward model.....	128
Figure 5.2. Plots of the first four principal components against time for three different tokens of the words ‘big’ (left) and ‘can’ (right).	136
Figure 5.3. Activation plots for the hidden unit number 15.....	137

Figure 5.4. Example representations for the words ‘big’ and ‘can’	137
Figure 5.5. Bar chart comparing the mean hidden plot distance of different words with the mean hidden plot distance of different tokens of the same word.....	138
Figure 5.6. Plots against time of a) the total amount of variance in hidden plot distance explained by the regression model, b) the amount of variance explained by phoneme similarity and c) the amount of variance explained by global similarity	140

Development of Phonological Representations in Young Children

Stephanie Ainsworth

The University of Manchester

**Doctor of Philosophy, School of Psychological Sciences, Faculty of Medical and
Human Sciences**

February 2015

Abstract

The development of phonological representations remains a hot topic within both the developmental and neural network literature. Historically, theoretical accounts have fallen within one of two camps: the accessibility account which proposes that phonological representations are adult-like from infancy (Rozin & Gleitman, 1977; Liberman, Shankweiler & Liberman, 1989) and the emergent account which proposes that phonological representations become gradually restructured over development (Metsala & Walley, 1998; Ventura, Kolinsky, Fernandes, Querido & Morais, 2007; Ziegler & Goswami, 2005). Within this thesis we tested predictions made by the accessibility account and key variants of the emergent account using data from both behavioural (Chapters 2, 3 and 4) and neural network studies (Chapter 5). The novel measures used within Chapters 2 to 4 were devised to allow us to contrast implicit measures of phonological representation (PR) which probe the segmentedness of the representations themselves, with explicit PR measures which tap into children's conscious awareness of phonological segments. Within Chapter 2 we present evidence that while explicit awareness of phonological structure is dependent on letter-sound knowledge, implicit sensitivity to the segments within words emerges independent of literacy. Within Chapter 3 a longitudinal study investigated the segmentedness of children's phonological representations at the rime and phoneme level. These results demonstrate that implicit sensitivity to both rime and phoneme segments is driven by vocabulary growth and is not dependent on letter-sound knowledge. The results within Chapter 3 also suggest that, while awareness of rime segments emerges naturally through oral language experience, explicit awareness of individual phonemes is related to letter-sound knowledge. In Chapter 4 we explored the idea of global versus phonemic representation using a mispronunciation reconstruction task. We found that sensitivity to both global and phonemic similarity increased over time, but with global sensitivity reaching adult levels early on in development. In Chapter 5 a neural network was trained on the mappings between real acoustic input and articulatory output data allowing us to simulate the development of phonological representations computationally. The simulation data provide further evidence of a developmental increase in sensitivity to both global and phonemic similarity within a preliterate model. Taken together, the results provide strong evidence that as children's vocabularies grow they become increasingly sensitive to both the global properties and segmental structure of words, independent of literacy experience. Children's explicit awareness of phonemes, on the other hand, seems to emerge as a consequence of learning the correspondence between letters and sounds. Within the context of the wider literature, the current results are most consistent with the PRIMIR framework which predicts early detailed phonetic representations alongside gradually emerging phonemic categories (Werker & Curtin, 2005). This thesis underlines the importance of using implicit measures when trying to probe the representations themselves rather than children's conscious awareness of them. The thesis also represents an important step towards modelling the emergence of segmental representation computationally using real speech data.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification at this or any other institute of learning.

Copyright Statement

The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see <http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487>), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see <http://www.manchester.ac.uk/library/aboutus/regulations>) and in The University’s policy on Presentation of Theses

Acknowledgements

I would like to thank all the children at Cadishead Primary School and Lily Lane Primary School who participated in my study. It was a real pleasure to work with you. Thank you also to the parents for providing consent for their children to take part. I would like to say a big thank you to the staff at both schools for all their help with recruitment, for finding me a space to conduct the study and for making me feel welcome. I would also like to thank Imogen Long, Amy Ogden and Louise Calland for their hard work in recruiting and testing the adult participants used as a control group within the thesis. Thank you to all my family and friends for their patience and support throughout my studies – I couldn't have done it without you! And last but by no means least I would like to thank my wonderful supervisors for all their advice, support and wisdom over the last few years. I have learnt an awful lot from each of you and have thoroughly enjoyed the experience.

This work was supported by the Economic and Social Research Council [ES/J500094/1].

The Author

I graduated from the University of Manchester Institute of Science and Technology (UMIST) in 2001 with an MPhys in Physics and Astrophysics. I then went on to study for the Postgraduate Certificate in Primary Education at Manchester Metropolitan University before beginning a career in teaching. I enjoyed seven years as a primary school teacher whilst also working towards the Graduate Diploma in Psychology part time at the University of Bolton. In 2010 I undertook a Masters of Research in Psychological Research Methods at the University of Manchester. My Masters research project was supervised by Dr Anne Hesketh and Dr Stephen Welbourne and investigated novel ways to measure the level of segmentation of children's phonological representations. Following completion of my Masters in 2011, I stayed on at the University of Manchester to undertake a three-year PhD. I am now working at Manchester Metropolitan University as a Senior Lecturer in Primary English Education.

CHAPTER 1

INTRODUCTION

Thesis Overview

This thesis is written in alternative format consisting of four journal articles, Chapters 2-5, preceded by an introductory chapter which outlines the key literature pertinent to the development of phonological representations and provides motivation for the research questions investigated within the four individual studies. A final chapter, Chapter 6, synthesises the research findings and discusses their limitations and implications for theories of phonological development, as well as suggestions for future work.

Phonological representations are the structures proposed to store information about the sounds in words. There is currently debate about how children's representations develop and how they interact with literacy experience. In particular there is disagreement about the relationship between the representations themselves, children's explicit knowledge of the phonological segments within those representations – often referred to as phonological awareness (or PA) – and children's knowledge of the mappings between letters and sounds (Walley, Metsala & Garlock, 2003; Ziegler & Goswami, 2005). Within this chapter we will outline the key theories of phonological development and discuss how they can be categorised in terms of two key themes: 1) accessibility versus emergence and 2) precipitation of the phoneme. We will also explore two key methodologies used to investigate the development of phonological representations: similarity judgement tasks and connectionist models.

Before we begin a review of current phonological theories it is important for us to clarify some key terminology which occurs regularly throughout the thesis. Phonological representations are the structures within which children store phonological information about words that they know. Phonological awareness on the other hand, is children's explicit knowledge of these structures, defined as children's ability to manipulate, analyse and reflect upon the phonological segments within words. Phonological awareness is often divided into two subtypes: rhyme awareness and phoneme awareness. While rhyme awareness refers to children's conscious awareness of rime segments (measured for example by asking children to say a word which rhymes with a stimulus word), phoneme awareness refers to children's conscious awareness of the individual phonemes within words (measured for example by asking children to segment a stimulus word into its constituent phonemes). Throughout the thesis we use the term implicit PR to describe measures, which are designed to probe the underlying segmentedness of children's phonological representations and *do not* require children to have a conscious awareness of

phonological segments (i.e. rimes and phonemes). The term explicit PR is used to describe measures of phonological representation, which *do* require a conscious awareness of phonological segments and therefore fall under the umbrella of phonological awareness measures.

Accessibility versus emergence: Are children's phonological representations adult-like from infancy?

There is longstanding debate within the literature about whether phonological representations develop gradually over time or whether they are adult-like from infancy. Within the latter view – often referred to as the accessibility position – it is claimed that infants already represent words at the same level of detail as adults, but that they can only access the phonological segments explicitly (e.g. when asked to break words up into phonemes in phoneme awareness tasks) later in childhood when they have developed the required meta-cognitive skills and alphabetic knowledge (Liberman & Shankweiler & Liberman, 1989; Morais, Bertelson, Cary, Alegria, 1986; Rozin & Gleitman, 1977). Within this view orthographic knowledge is seen as the key to unlocking explicit awareness of the sound structure of words which have been stored within children's phonological representations since infancy. Evidence in support of the accessibility account comes from studies which indicate early phonological specificity (Bailey & Plunkett, 2002, Ballem & Plunkett, 2005, Swingley & Aslin, 2000; Swingley & Aslin, 2002; Swingley, 2009, see also Ramon-Casas & Bosch, 2014 for a review) and studies showing a steep rise in phoneme awareness (explicit awareness of the individual phoneme within words) at the onset of literacy instruction (e.g. Duncan, Cole, Seymour, & Magnan, 2006; Mann & Wimmer, 2002; see Ziegler & Goswami, 2005 for a review). The opposing view, often referred to as the emergent account (Walley, Metsala & Garlock, 2003), argues that infants' phonological representations are qualitatively different from adults' and are gradually restructured driven by oral language (Metsala & Walley, 1998; Ventura, Kolinsky, Fernandes, Querido & Morais, 2007) and/or literacy experience (Ziegler & Goswami, 2005).

One variant of the emergent account – the lexical restructuring model (Metsala & Walley, 1998) – proposes that when young children only know a limited number of words they do not need to store them in a lot of detail in order to keep the representations of different words distinct. As they learn more words however, their phonological representations need to be restructured to avoid confusion between similar sounding words. Within the lexical restructuring model (LRM) children's phonological representations

initially have a global/holistic form characterised by the overall acoustic shape of a word or by one particularly salient feature. As children's vocabularies grow, their representations are proposed to become increasingly segmented into onset-rime (e.g. cat stored as /k/-/æt/) and then finally into phonemic form (e.g. cat stored as /k/-/æ/-/t/). Lexical restructuring is proposed to take place gradually over childhood with high frequency words from dense neighbourhoods being restructured first. Evidence in support of the LRM comes from studies showing an interaction between frequency and neighbourhood density on a range of speech processing tasks (Metsala, 1999) and studies which show qualitative differences between the way that adults and children process words (Carroll & Myers, 2011; Cole & Perfetti, 1980; Walley, 1987; Walley & Metsala, 1990; 1992; Treiman & Breaux, 1982). For example on a mispronunciation detection task, while adults were found to identify mispronunciations more easily in initial position, four year old children did not show any effect of position (Walley, 1987). These results are interpreted by Metsala and Walley (1998) as evidence that adults' representations are more segmental than children's.

Further support for the emergent account comes from Fikkert (2010) who found evidence that infants aged 14 months have global representations which are underspecified in terms of place of articulation. The results suggest that the place of articulation for the whole word is represented as the place of articulation of the stressed vowel within the word. Other work by Fikkert and colleagues provides further evidence of early representations being underspecified (Altvater-Mackensen, van der Feest & Fikkert, 2013), with 18 month olds found to detect changes from stop to fricative but not fricative to stop. Here the results are interpreted in terms of a framework where early representations are specified by 'marked' features only and become augmented with other featural information over development.

Conversely, several infant studies provide evidence that children's phonological representations are fully specified from infancy (Bailey & Plunkett, 2002, Ballem & Plunkett, 2005, Swingley & Aslin, 2000; Swingley & Aslin, 2002; Swingley, 2009; Ramon-Casas & Bosch, 2014). For example an influential study conducted by Bailey and Plunkett (2002) found both 18 month and 24 month old toddlers to be sensitive to mispronunciations involving minimal contrasts (words which had been changed by only one or two features) suggesting that infants representations already contain fine levels of detail. Moreover, performance was found not to be related to age, vocabulary size, the neighbourhood density or age of acquisition of the words, in contention with the predictions made by the lexical restructuring model.

Numerous similar studies followed, providing further evidence of detailed early phonological representations (see Ramon-Casas & Bosch, 2014 for a review), with

children as early as 14 months also found to be sensitive to minimal mispronunciations (Ballem & Plunkett, 2005). Proponents of the emergent view however, have pointed out that many of these studies have tested infants using mispronunciations in initial position only and have suggested that the results could be explained in terms of the salience of initial parts of words without requiring infants' representations to be segmental (Bowey & Hirakis, 2004). More recently however, Swingley (2009) reported evidence that infants are sensitive to mispronunciations of codas as well as onsets. Again in line with Bailey and Plunkett's study (2002), no relationship between vocabulary size and mispronunciation effect was found (Swingley, 2009).

A counter argument made in response to claims of early adult-like representations is that sensitivity to phones does not necessitate phonemic representation (Ziegler & Goswami, 2005). In other words, the fact that infants are sensitive to small changes in words does not mean that their phonological representations are necessarily segmental. Ziegler and Goswami (2005) point out that the ability to recognise that /ba/ is different from /pa/ does not require the /p/ in 'pit' to be stored in the same way as the /p/ in 'lap', which will differ in its exact phonetic manifestation. While the former ability requires sensitivity to phones, the latter ability requires abstract representation of word forms at the phoneme level. In order to investigate how segmental children's phonological representations are, we therefore need to use measures which probe representation of phonemes (categories abstracted across different words) rather than measures which detect sensitivity to phones (utterances of a given phoneme within a particular word).

In the commentary above we have summarised evidence for early detailed phonological representations on the one hand, and representations which mature gradually over development on the other. However, as highlighted by Swingley (2009), if we reframe our definition of the term 'lexical restructuring', the two accounts needn't be contradictory. One way to marry evidence in support of lexical restructuring with evidence of early phonological specificity is to assume that children's phonological representations contain fine levels of detail from the outset but that children's interpretation of the phonetic variability within them changes over development (Swingley, 2009). According to Swingley's hypothesis, as children's vocabularies grow they gradually abstract phonemic categories from the statistical regularities across their stored (phonetically detailed) representations for different words. In contrast to the LRM's idea of early underspecified phonological representations becoming increasingly segmental over time, Swingley (2009) frames restructuring in terms of phonemic probability distributions (present from infancy) being gradually sharpened over childhood. Evidence in support of Swingley's account comes from studies which indicate that although infants can detect small changes in words

(Bailey & Plunkett, 2002, Ballem & Plunkett, 2005, Swingley & Aslin, 2000; Swingley & Aslin, 2002; Swingley, 2009; Ramon-Casas & Bosch, 2014) they do not appear to interpret deviant pronunciations as new words (Merriman & Schuster, 1991; Stager & Werker, 1997; Swingley & Aslin, 2007; White & Morgan, 2008). Swingley (2009) explains these findings in terms of the correct pronunciation having a higher probability than the mispronunciation, but with the deviant form remaining as a possible, although less probable, variant of the target word. Further support for Swingley's account comes from his earlier study (Swingley, 2007) which suggested that phonemic categorisation varies as a function of lexical exposure and more generally from studies which indicate significant changes in categorisation beyond infancy (Hazan & Barrett, 2000; Nittrouer, 2002; Walley & Flege, 1999).

The notion of early detailed representation alongside the gradual emergence of phonemes is also embodied within the PRIMIR framework: Processing Rich Information from Multidimensional Interactive Representations (Werker & Curtin, 2005). Within PRIMIR early words are stored in rich detail from infancy in the form of word level exemplars which cluster within the word form plane. These exemplars are proposed to contain both contextual and indexical information, i.e. about the way a word sounds within a particular sentence (contextual) when spoken by a particular speaker (indexical). Evidence in support of rich contextual information being stored from infancy comes from empirical studies which show that the representations of one year-old infants contain non-criterial information, i.e. detail that is not needed to contrast two different words (Curtin, 2011). PRIMIR proposes that as infants' vocabularies grow, higher order regularities begin to emerge within the phoneme plane, reflecting the beginnings of phonemic categories. Although 'phonemes' start to emerge within the phoneme plane during infancy, they only become adult-like after being 'sharpened up' through the process of learning to read (Werker & Curtin, 2005). Within PRIMIR rich information from the speech signal is stored across multidimensional planes (general perceptual, word form and phoneme plane) throughout development, but access to the different levels of information is filtered according to the developmental level of the child and the demands of the task at hand. In this way PRIMIR is therefore able to account for a number of seemingly contradictory findings within the infant literature (Werker & Curtin, 2005). For example, while infants aged 14 months have been shown to be sensitive to fine levels of phonetic detail (Swingley & Aslin, 2002), they confuse minimally contrasting items during word learning (Werker, Fennell, Corcoran & Stager, 2002).

Within both Swingley's account (Swingley, 2009) and the PRIMIR framework (Werker & Curtin, 2005) children's representations are proposed to contain rich contextual

detail from the outset, with phonemic categories gradually being extracted from the regularities between similar sounding words. This is in contrast with the lexical restructuring model (Metsala & Walley, 1998) which assumes that young children's phonological representations are initially lacking in both detail and segmentedness (Metsala & Walley, 1998). Within the LRM detail is added as phonological representations undergo a whole to parts shift from holistic/global representations to onset-rime and then finally phonemic representation. We can therefore think of restructuring in two different ways: 1) in terms of probability distributions gradually sharpening into phonemic categories and 2) in terms of global representations becoming segmented into phonemic segments (possibly with an intermediate period of onset-rime representation).

In summary, there is debate within the literature about how detailed infants' phonological representations are, and if and how they change over development. While the majority of studies indicate that infants are sensitive to very small changes in pronunciation (Ramon-Casas & Bosch, 2014), it is not clear whether or not this reflects early phonological representations which are qualitatively similar or different from those of adults. While numerous studies have provided evidence of early sensitivity to changes at the phone level, fewer studies have investigated emergence of phonemic representation. Within Chapter 2 this issue is investigated using novel implicit measures which probe the level of segmentedness of children's phonological representations at the phoneme level. This study tests the prediction made by the LRM that children will be less sensitive to shared segments than adults and that segmental sensitivity will increase over development (Metsala & Walley, 1998). This work is extended within Chapter 3 with a longitudinal study tracking children's segmental sensitivity and vocabulary growth over the first two years of school. Within this chapter children's performance is analysed at different grain sizes allowing us to test theoretical predictions about the development of onset-rime versus phoneme level representation. Chapter 4 investigates the shift from global to phonemic representation predicted by the LRM (Metsala & Walley, 1998) by tracking sensitivity to both global and phonemic similarity relations using cross-sectional and longitudinal data. Within Chapter 5 the theme of accessibility versus emergence is investigated using a neural network simulation of the development of phonological representations. Within this study a model is trained on the mappings between real acoustic and articulatory data. The model's hidden unit representations are then analysed to assess whether they become increasingly segmental over training and to investigate how the model's sensitivity to global similarity changes over time.

Precipitation of the phoneme: Does phonemic representation emerge naturally in the absence of literacy?

Another area of contention within the literature relates to whether preliterate children represent words in terms of phonemes or whether it is only once they learn about phonemes through literacy experience that phonemic representation begins to appear. Several studies have shown that orthographic knowledge affects phonological processing (e.g. Ziegler & Ferrand, 1998; Seidenberg & Tanenhaus, 1979; Muneaux & Ziegler, 2004). For example in Frith's pioneering imaging study (1998) adults who were either literate or illiterate showed different patterns of activation when repeating nonwords, with illiterate adults showing greater activation in the pre-frontal cortex than literate adults. Frith (1998) proposed that while illiterate adults process nonwords as new semantic items, literate adults process nonwords phonologically by automatically segmenting them into phonemes. Frith went on to compare the effect of orthographic knowledge on speech processing as being akin to catching a virus: once letter-sound mappings have been learnt 'language is never the same again' (Frith, 1998, p.1011) and heard words are henceforth automatically segmented into chunks of sound.

Muneaux and Ziegler (2004) have also proposed that orthographic knowledge increases the level of segmentation of children's phonological representations. Muneaux and Ziegler (2004) found evidence that on a neighbour generation task (where adults were asked to think of words differing by one phoneme from the stimulus) participants produced more phonographic neighbours (words which also differ by only one grapheme) than phonological neighbours (words which differ by one phoneme but are not phonographic neighbours). This finding was interpreted as evidence that orthographic knowledge leads to further restructuring of phonological representations, in addition to the changes in representation driven by vocabulary growth proposed within the LRM. However, the apparent effects of orthography on phonological representations observed by Muneaux and Ziegler (2004) and in other studies (Ziegler & Ferrand, 1998; Seidenberg & Tanenhaus, 1979) could also be explained in terms of coactivation of separate orthographic representations and phonological representations (Grainger & Ferrand, 1994) rather than in terms of orthography actually changing the phonological representations themselves. Pattamadilok and colleagues (Pattamadilok, Knierim, Duncan & Devlin, 2010) set out to test which of the two possibilities is behind the observed orthographic effects on speech processing using transcranial magnetic stimulation (TMS). Within this study orthographic consistency effects (speeded recognition for words whose rime can only be spelt one way, Ziegler & Ferrand, 1998) were measured during TMS stimulation to the left supramarginal gyrus (SMG), an area usually associated with phonological processing, and the left ventral

occipitotemporal cortex (vOTC), normally associated with orthographic processing. The results provide evidence in support of the first possibility with orthographic consistency effects found to disappear when TMS stimulation was used to disrupt the phonological processing area (SMG) but remained when the orthographic area (vOTC) was disrupted (Pattamadilok et al., 2010). Pattamadilok et al. (2010) therefore concluded that orthographic knowledge leads to direct restructuring of phonological representations while acknowledging the possibility that coactivation of visual representations may also have an additional effect on phonological processing. A key implication of Pattamadilok et al.'s study is that phonological representations only become segmental once children learn the links between letters and sounds.

Pattamadilok et al.'s findings (Pattamadilok et al., 2010) echo the earlier proposition made within psycholinguistic grain size theory (Ziegler & Goswami, 2005) that orthographic knowledge is a prerequisite for the emergence of the phoneme within the lexicon. Ziegler and Goswami (2005) argue that it is only once children learn about letters that they become aware of phonemes and consequently begin to organise their lexicon phonemically. Further support for this hypothesis, in addition to the brain imaging work described above (Pattamadilok, Knierim, Duncan & Devlin, 2010), comes from studies showing a rapid rise in phoneme awareness at the onset of literacy instruction (e.g. Ehri, 1979; Morais, Cary, Alegria, & Bertelson, 1979; Read, Zhang, Nie, & Ding, 1986, see Ziegler & Goswami, 2005 for a review). However, while Ziegler and Goswami (2005) propose that orthography precipitates the emergence of phonemes within the representations themselves, the supporting evidence which they draw upon is based upon studies of explicit phoneme awareness rather than upon studies investigating the underlying representations. It is therefore possible that phonemes emerge within the lexicon before children are able to access them explicitly on traditional phoneme awareness tasks. Within the current thesis (Chapters 2 and 3) we aim to address this issue by contrasting the measurement of children's implicit phonological knowledge (referred to as implicit PR within Chapters 2 and 3) with performance on traditional PA measures which measure explicit phonological awareness (referred to as explicit PR measures within Chapters 2 and 3). By using these measures within a sample of preschool children who have varying levels of literacy experience, we are able to investigate the extent to which segmental phonology at the representational level emerges independent of letter-sound knowledge.

PRIMIR also proposes that literacy experience plays a key role in the development of segmental phonological representation (Werker & Curtin, 2005); however (as mentioned above) within the PRIMIR framework, the emergence of phoneme-like

representations is proposed to begin before children learn about letters. Letter-sound knowledge then serves to ‘sharpen up’ these categories as children learn to read. PRIMIR therefore suggests that both vocabulary growth and letter-sound knowledge are key drivers of segmental representation, but that children’s phonemic categories do not reflect those of adults until they learn they learn the links between letters and sounds (Werker & Curtin, 2005).

Not all researchers however agree that orthographic knowledge is central to the development of phonemic representation. The key driver of representational change within the lexical restructuring model (LRM) is vocabulary growth which is proposed to force a transition from global, holistic representation to phonological representations which are segmented into onsets and rimes then finally into individual phonemes (Metsala & Walley, 1998). While the LRM does not make any specific predictions about the role of letter-sound knowledge it is implied that the arrival of many phonologically close neighbours is sufficient to restructure words so that they are segmented into phonemes. Evidence in support of lexical restructuring independent of literacy comes from a study which contrasted the performance of literate and illiterate adults on a number of phonological processing measures (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007). Ventura and colleagues found evidence of an interaction between neighbourhood density and frequency within both literate and illiterate adults. Given that such an interaction has been proposed to indicate segmental restructuring (Metsala & Walley, 1998), Ventura et al. (2007) concluded that lexical restructuring occurs without the need for significant orthographic knowledge.

Ainsworth and colleagues have also found evidence consistent with the idea of restructuring in the absence of literacy, within a sample of nursery children aged 3;6 to 4;6 (Ainsworth, Welbourne & Hesketh, submitted). Within this study a set of novel implicit PR measures were developed to probe the phonological representations of preliterate children, without requiring them to have any explicit awareness of phonological structure (e.g. of the chunks of sound within words). Ainsworth et al. (submitted) found that nursery children who had not yet begun formal literacy instruction showed sensitivity to the number of shared phonemes between words, suggesting that literacy may not be required for the emergence of segmental representation. There were however two key limitations to this work: firstly the study did not include a measure of literacy knowledge and so we cannot be sure of how many children were in fact preliterate; secondly there was a potential confound of global similarity which may have acted to inflate children’s apparent segmental sensitivity. The issue of global similarity is discussed in detail below. In Chapters 2 to 4 we overcome these issues by using similar measures of implicit

phonological knowledge which have been adapted to allow us to control for global similarity (see below) alongside a direct measure of letter-sound knowledge.

In the above commentary we have seen that there is disagreement within the literature as to whether or not literacy experience is needed before children's phonological representations become phonemic. There is also debate about the relationship between the segmentedness of children's phonological representations and children's explicit phonological awareness. Within psycholinguistic grain size theory the effect of letter-sound knowledge on phonological representations is mediated by phoneme awareness. The idea here is that when children are taught about the correspondences between letters and sounds (or more generally between graphemes and phonemes) this sparks an awareness that words can be analysed in terms of phoneme categories. Ziegler & Goswami (2005) suggest that it is only when children gain this awareness of phonemes that children start to represent words in that way. Within the LRM however, the direction of causality is reversed with the emergence of phonemes at a representational level setting the stage for phoneme awareness. Here the idea is that we first need to have the phonemic structure within our representations before we can gain explicit access to the phonemes within words (Metsala & Walley, 1998). Consistent with the LRM's proposition, Caravolas and Landerl (2010) found evidence that phoneme awareness is influenced by oral language experience prior to the development of orthographic knowledge. Cross linguistic differences in patterns of performance on phoneme awareness tasks across samples of Czech and German samples were found in pre-reading children who had very little literacy experience. Caravolas and Landerl (2010) interpret their results as evidence that letter-sound knowledge is not a prerequisite for phoneme awareness, rather phoneme awareness is firmly rooted in early language experience.

Within the current work the relative roles of oral language and literacy experience on the development of children's phonological representations were investigated by measuring children's phonological knowledge alongside both vocabulary growth and letter-sound knowledge (Chapters 1 and 2). Inclusion of both implicit and explicit measures of phonological representation (PR) allowed us to probe the segmentedness of children's representations on the one hand (using implicit PR measures), with children's conscious awareness of phonological segments on the other (using explicit PR measures – often referred to elsewhere as phonological awareness or PA tasks, see below). This in turn allowed us investigate whether letter-sound knowledge is needed before children represent words phonemically and/or become explicitly aware of phonemes. Within Chapter 4 phonological development of phonological representations was simulated within a connectionist computational model. The networks were analysed over time allowing us to

investigate whether sensitivity to phonemes emerges through oral language experience alone within a model which has no orthographic knowledge.

Note that while measures of explicit knowledge of the sounds within words are traditionally referred to as phonological awareness tasks, within the current work we will use the term explicit PR measures. This is to emphasize the contrast between segmental phonology at the representational level (as measured by implicit PR tasks) and children's explicit knowledge of the phonological segments within words (as measured by explicit PR tasks).

Similarity judgements as a probe into the lexicon

In order to investigate the predictions made by different theoretical accounts of phonological development, researchers have attempted to probe the structure of children's phonological representations. One way to do this is to ask children to make similarity judgements about phonological word forms, the assumption being that if children judge two phonological forms sharing a given set of properties as similar, this indicates that they represent words in terms of these properties (Carroll & Myers, 2011; Storkel, 2002; Treiman & Breaux, 1982). For example, Storkel (2002) asked children to listen to spoken words and decide whether or not they sounded like another word (e.g., 'Is /tʌg/ like /tʌf/?'). By manipulating the similarity relations between the items, Storkel was able to investigate whether children store words in terms of shared phonemes, shared manner or shared place of articulation. Storkel's results suggest that while children store words from dense neighbourhoods in terms of shared phonemes, they represent words from sparse neighbourhoods more coarsely, with the final sound represented in terms of shared manner (e.g. 'boom' stored as /b/-/u/-/nasal/. Carroll and Myers (2011) used a similar set of judgement measures to investigate developmental changes in phonological representations. While both children and adults made manner classifications (i.e. judging words with shared manner of articulation as similar) as well as phonemic classifications, phonemic classifications became more frequent with age and manner classifications tended to be in final position, consistent with Storkel's findings (Storkel, 2002).

Similarity judgement tasks have tended to use a forced choice format requiring children to say whether a pair of words is similar or not (Carroll & Myers, 2011; Storkel, 2002) with one notable exception (Treiman & Breaux, 1982 discussed below). Multiple choice is potentially more powerful than forced choice when trying to distinguish whether or not children are using phonemic similarity relations. This is because in a forced choice task there is no set threshold for whether something is similar or not – rather similarity is relative. For example, a literate adult who presumably represents words at the phoneme

level might consider the words 'beach' and 'dish' to be similar based on global features, e.g. they both start with a stop and end with an affricate/fricative (Carroll & Snowling, 2001). But when asked to choose which of two words is most similar to 'beach' – 'bean' or 'dish', they might judge 'bean' to be more similar on the basis that it has two shared phonemes, whereas 'dish' has none. In other words, while a forced choice task tells us whether two items are classed as similar or not, it does not tell us directly about graded similarity, i.e. which type of similarity relation is considered to be most important.

Treiman and Breaux's study (1982) adopted a multiple choice format, by asking children to choose the most similar pair of pseudowords from a group of three (e.g. 'Which two words sound the most alike: /bis/, /diz/ or /bun/?'). The items were constructed so that there would be a common phoneme pair (/bis/ and /bun/), an overall (global) similarity pair (/bis/ and /diz/) and an anomalous pair (/diz/ and /bun/) which are dissimilar both globally and phonemically. Children were found to make judgements based on global rather than phonemic similarity, whereas adults showed the reverse pattern. This was interpreted as evidence that children predominantly use global relations when making similarity judgements, whereas adults predominantly use phonemic relations. The results need to be interpreted with caution however, given that neither the child or adult difference between global and phonemic response reached significance (children made global judgements 44% of the time and phonemic judgements 37% of the time; the corresponding percentages for adults were 43% and 48%). The fact that adults chose the phonemic match less than half the time suggests that either adults do not base similarity judgements on phonemes either, or that the task is not sufficiently sensitive to detect these relations reliably. A possible reason for the apparent lack of sensitivity of this task is that the working memory demands are too high. The task involves holding three pseudowords in memory while also comparing them with one another pairwise before making a judgement.

Ainsworth and colleagues (Ainsworth, Welbourne & Hesketh, submitted) constructed a novel similarity based judgement task which reduces the demands on short term memory by using picture stimuli. Within this task (named mispronunciation reconstruction) an alien mispronounces a word and the child guesses which of four words the alien is trying to say. By presenting pictures of the four words as the child hears the mispronunciation, the task was designed so that only one token (the mispronunciation) needed to be held in memory. The number of shared phonemes between the mispronunciation and the four pictured items was manipulated so that the authors could look for evidence of segmental representation. The study found evidence of sensitivity to the number of shared phonemes within a sample of nursery children (aged 3;6 to 4;6) and illustrated the feasibility of using similarity based judgement tasks with preschool children.

While the findings of the pilot study are consistent with the idea of phoneme level representation preceding formal literacy exposure (Ainsworth, Welbourne & Hesketh, submitted), it is possible that phonemic sensitivity was confounded by global similarity as discussed below.

Global similarity relates to how similar two words are when compared holistically, in contrast to phonemic similarity which relates to how similar two words are in terms of the number of shared phonemes. The global similarity of distracters has been shown to affect children's performance in phoneme invariance (Byrne & Fielding-Barnsley, 1993) and rhyme and alliteration matching tasks (Carroll & Snowling, 2001). For example, Byrne and Fielding-Barnsley (1993) found that children's performance levels fell when the distracters were matched for global similarity, with half of the children who passed the original task failing once global similarity was controlled for. Before distracters were matched some of the children appeared to be comparing words on a phoneme by phoneme basis, e.g. choosing 'bowl' over 'shed' as the word which starts with the same sound as 'beak' but were actually choosing 'bowl' because it was the most similar globally to 'beak' (Byrne & Fielding-Barnsley, 1993). Both studies concluded that unless items are effectively matched for global similarity, task performance may provide an overestimation of the child's true phonological awareness level (Byrne & Fielding-Barnsley, 1993; Carroll & Snowling, 2001). The same argument also applies to multiple choice similarity tasks designed to probe the segmentedness of the representations themselves. For example, a child may choose 'bed' as sounding more similar to 'bin' than 'hat', not because they are making a similarity judgement based on phonemic representation, but because 'bed' sounds more like 'bin' than 'hat' overall. Controlling for global similarity in both measures of phonological representation and phonological awareness is therefore very important. Within Chapters 1 and 2 a set of novel similarity judgement measures is used which have been designed to probe sensitivity to the number of shared phonemes while controlling for global similarity.

Although global similarity has been identified as an important potential confound within similarity based paradigms, researchers have yet to investigate sensitivity to global similarity as a competency in its own right. As discussed earlier in this chapter, the idea of 'global' representation features explicitly within the lexical restructuring model (Metsala & Walley, 1998). Although Metsala and Walley (1998) make no direct predictions about how children's sensitivity to global similarity might change over time, we might expect that the shift from global to segmental phonological representation would lead to a developmental fall in global sensitivity alongside a rise in phonemic sensitivity. The concept of global similarity is also relevant to (although not referred to within) PRIMIR (Werker & Curtin,

2005), which proposes that both a word level plane and phoneme plane operate throughout development. We might expect therefore within PRIMIR that sensitivity to global similarity relations will remain important, alongside an emerging sensitivity to phonemes. Within Chapter 3 one of the novel measures of phonological representation (mispronunciation reconstruction) is used to separate the two sensitivities and to plot how they change over development.

A connectionist approach to the development of phonological representations

We have seen above how similarity judgement tasks can be used to investigate the development of phonological representations. An alternative approach is to model phonological development computationally. As highlighted by Harm, McCandliss and Seidenberg (2003), computational modelling has a number of advantages: Modelling does not involve taking children out of lessons, negotiating access to schools or gaining permission from parents. It also enables the mechanistic testing of causal hypotheses and allows us to implement extreme conditions (e.g. the absence of any literacy experience) which would be unethical or impossible to impose on children (Mareschal & Thomas, 2007). While space allows only a brief review, we will outline a few of the most relevant modelling studies to phonological development, focussing on how they have attempted to simulate phonological representation.

Connectionist models of reading

Seidenberg and McClelland (1989) developed a connectionist model investigating how people learn to read aloud. Seidenberg and McClelland (1989) simulated the model's phonological representations with Wickelphones (Wickelgren, 1969) – representations corresponding to a central phoneme and its two closest neighbours – using the system developed by Rumelhart and McClelland (1986). Within this system the word 'soft' is represented by the Wickelphones [_so], [sof], [oft] and [ft_] where '_' represents the beginning or end of a word. Each Wickelphone activates 16 Wickelfeatures which store phonetic information about each of the three phonemes along four articulatory dimensions. While the first two dimensions act to distinguish vowels from consonants and to coarsely divide phonemes according to manner, the second dimension subdivides each manner class and separates high from low vowels, the third dimension codes the place of articulation and the fourth codes phonemes in terms of voicing (for consonants) and duration (for vowels).

The rationale behind the Wickelfeature system is that it allows learning about a phoneme in one context to be generalized to other contexts. Using 'soft' again as an example, the learning embedded within the connections for the Wickelphone [oft] is accessible not only when activated by the word 'soft' but also when activated by other words sharing the same Wickelphone such as 'loft' and 'often'. By representing words in a way that produces shared activation between phonologically similar sounding words, the model allows learning about a particular sound pattern in one word to be generalized to other words. Seidenberg and McClelland's (1989) model was able to account for why some words are harder to process than others, why some people are better than others at recognising written words, how reading in silence differs from reading aloud and for some aspects of dyslexia (Seidenberg & McClelland, 1989). Not all aspects of the model fitted well with behavioural observations however, with poorer performance than expected for the pronunciation of nonwords.

Plaut, McClelland, Seidenberg and Patterson (1996) attempted to eliminate the discrepancies between simulated and real nonword performance found by Seidenberg and McClelland (1989) by using a finer grained system for the orthographic and phonological representations. In the new system phonological representations were represented by phonemes rather than by Wickelfeatures, with every letter activating all of its corresponding phonemes, thus allowing generalization. Previously this type of representational scheme had not been used because in doing so positional information would be lost, e.g. 'pat' and 'tap' would be indistinguishable (Plaut et al., 1996). However Plaut et al. (1996) noted that if you incorporate phonotactic constraints into the coding the spelling-sound correspondences can be 'condensed' (Plaut et al, 1996, p.11) and both generalization of learning and positional information can be maintained. In other words representation became easier to do on a phoneme by phoneme basis once it was taken into account that (with only a few exceptions) consonants within an initial or final cluster can only take one position (e.g. if /b/ and /l/ are within an initial cluster there is only one allowed order for them). It was found that when this improved system for phonological representation was implemented the model produced the expected performance for nonwords (Plaut et al., 1996).

Harm and Seidenberg (1999) then extended the two previous models (Seidenberg & McClelland, 1989; Plaut, McClelland, Seidenberg and Patterson, 1996) by modelling how phonological representations are stored before children learn to read, and the reciprocal relationship between phonological representations and reading once instruction begins. Within this model phonological representations were simulated using phonemic output slots within an attractor network. In the pre-reading phase the network was successfully

trained to represent phonological word forms and to ‘fill in’ missing parts of words via attractor basins (sets of states which the model tends towards). Once the network had learnt to read, the attractor networks were found to get stronger, the knowledge of phonotactic constraints grew, the network became more sensitive to rime (e.g. storing ‘eat’ in a similar way to ‘treat’ and ‘meat’) and became better at completing phonological patterns. From these findings the authors concluded that the introduction of orthographic knowledge cause the model’s representations to become more segmental (Harm & Seidenberg, 1999).

In a later study Harm, McCandliss and Seidenberg (2003) used their reading model to simulate the effect of reading interventions on impaired networks. In particular Harm and colleagues (2003) were interested in whether phoneme awareness training should be oral only (e.g. blending and segmenting phonemes out loud) or whether letter-sound correspondences should also be included. Interventions were implemented at three different stages: just as reading instruction begins, after 10,000 words and after 100,000 words have been presented. Two interventions were simulated: a phonological intervention which involved repairing the lost connections and stopping the weights from decaying, and a Word Building intervention (McCandliss, Beck, Sandak & Perfetti, 2003) which focussed on letter-sound correspondences. While the phonological intervention was only successful when implemented at the onset of reading instruction, the Word Building intervention improved reading performance at all three stages. The Word Building technique was shown to bring the phonological representations of similar words closer together and to increase the separation of dissimilar words. Within Harm, McCandliss and Seidenberg’s model (2003) phonological representations were modelled in the same way as in the Harm and Seidenberg (1999) model.

Connectionist models of spoken language

While the studies discussed so far have focussed primarily on reading, other connectionist models have focussed on the acquisition of spoken language (McClelland & Elman, 1986; Plaut & Kello, 1999; Westermann & Miranda, 2004; also see MacWhinney, 2010 for a review of recent corpus based studies). McClelland and Elman’s (1986) TRACE model of spoken word recognition simulates the learning of phonological forms with items stored across three layers: features, phonemes and whole words. Within this model the network’s phonological representations consist of mock phonemic speech inputs (a label assigned to each phoneme) and feature vectors (calculated using an abstract system based on a number of theoretically derived articulatory dimensions). TRACE successfully

reproduced several key empirical phenomena. For example, the model was able to simulate both top down (word level) and bottom up (feature level) influences on the identification of spoken phonemes. The model does however have a number of important limitations as noted by the authors (McClelland & Elman, 1986). There are two limitations which are especially relevant to the current thesis. Firstly, the model has no centrally stored representation and therefore requires repeated copying of units over the three levels. This limits both the efficiency and ecological validity of the model. Secondly, the model uses artificial inputs which impose a segmental structure onto the network as discussed in more detail at the end of this section.

Plaut and Kello (1999) investigated the formation of PRs by training a network in babbling, comprehension, imitation and intentional naming. Within this model articulatory representations were coded in terms of the following six dimensions: oral constriction, nasal constriction, place of constriction, tongue height, tongue backness and voicing. Acoustic representations were coded in terms of ten dimensions based on the formant frequencies, formant transitions, frication, burst, loudness and jaw openness. Past and current articulatory states were then mapped onto acoustic states using ten equations. A key property of Plaut and Kello's model (1999) is that phonology is represented as a hidden layer, whose structure is not predefined. This is theoretically important because it means that the architecture of the network is not constraining the way that phonological representations should be structured – the model is not making any assumptions about the 'natural' units of speech (Plaut & Kello, 1999). This is in contrast to the reading models and also the TRACE model discussed above where phonology is represented overtly within a predefined processing unit. The rationale behind the design of Plaut & Kello's model (1999) was to allow phonological representations to be learned (rather than pre-specified), driven by the functions of understanding and producing speech. Their results suggest that phonological representations become increasingly segmental through experience of spoken language alone. The model's errors were analysed at each developmental stage and phonemic errors (errors which are phonemically similar to the target) became more frequent over training.

Kello and Plaut (2004) point out however that their findings (from Plaut & Kello, 1999) may be, at least in part, an artefact of the model itself. Although the use of hidden units partly avoids imposing any predefined structure onto the way that words are stored, the segmental nature of the input parameters could be acting to impose phonemic structure onto the model's phonological representations. In order to establish whether segmental sensitivity does emerge naturally through language experience, we need a model whose design does not bias it towards segmental representations (Kello & Plaut 2004). Kello and

Plaut (2004) took a first step towards addressing this problem by using real speech instead of theoretically derived parameters to successfully train a feed-forward model on the mappings between articulatory and acoustic data. While Kello and Plaut's (2004) study proved the feasibility of using real speech data to train a connectionist network, the authors did not investigate the nature of the network's representations or how they changed as training progressed. Within Chapter 5 we build on Kello and Plaut's work (2004) by training a model on real speech data and analysing the nature of the model's hidden unit representations in search of evidence of emerging segmental structure. By using real speech data and an architecture which is as theoretically neutral as possible, we are able to test hypotheses about the precipitation of the phoneme without introducing an a priori bias towards phonemic structure into the model.

Research Aims

The overarching aim of this thesis is to investigate the development of children's phonological representations. As outlined above, current theories make differing predictions in relation to two key themes: 1) accessibility versus emergence and 2) the precipitation of the phoneme. In this thesis we test these predictions using both behavioural and simulation data. The first specific aim of the thesis was to investigate whether children's phonological representations become increasingly segmental over development as predicted by variants of the emergent account. To achieve this aim a battery of novel implicit tasks was used to measure sensitivity to the shared segments between words with both children (aged 3;2 to 5;7) and adults (Chapter 2). Measures of vocabulary size and letter-sound knowledge were taken to allow us to assess whether any emerging segmental sensitivity was driven predominantly by vocabulary growth or letter-sound knowledge. Explicit measures of phonological representation were also included to allow us to contrast children's implicit segmental sensitivity with their conscious knowledge of the sound segments within words (phonological awareness).

Within Chapter 3 this work was then extended in a longitudinal study using the same measures, which set out to establish whether phonemic representation emerges in the absence of literacy or whether it emerges through oral language experience alone. The longitudinal study provided us with the power to analyse children's segmental sensitivity at both the onset-rime and phoneme level. In doing this we were able to check whether any emerging segmental sensitivity in preliterate children could be accounted for solely by sensitivity to onset-rime segments or whether there is evidence of sensitivity to phonemes before children learn about letters.

Within Chapters 4 and 5 we aimed to separate children's sensitivity to phonemic versus global similarity and investigate how they both change over time. Although previous studies have considered the confounding influence of global similarity on phonological awareness tasks, they have not investigated global sensitivity as an important competency in its own right. Given that the idea of global representation is relevant to both the lexical restructuring model (Metsala & Walley, 1998) and PRIMIR (Werker & Curtin, 2005) it is important for both types of sensitivities to be investigated. Within Chapter 4 we analysed performance on a novel mispronunciation reconstruction task where both phonemic and global similarity relations had been manipulated. By analysing cross-sectional and longitudinal patterns of performance on this task we set out to separate children's sensitivity to phonemic versus global similarity and track them over development. Within Chapter 5 we addressed the same aim by investigating how a computational model's sensitivity to the two types of similarity changes as it is exposed to speech sounds and learns the motor gestures that produce these sounds.

Acknowledgement of Contribution of Other Authors

Dr Anne Hesketh, Dr Stephen Welbourne and Dr Anna Woollams supervised the work within this thesis. Dr Anna Theakston acted as project advisor. I devised and created all the behavioural tasks used within Chapters 2, 3 and 4. I also conducted all recruitment and behavioural testing of children for the work presented in this thesis. The testing of the adult control group within Chapter 2 was conducted by three final year Psychology students: Amy Ogden, Imogen Long and Louise Calland under the supervision of myself, Dr Anne Hesketh and Dr Stephen Welbourne. I conducted all of the computational modelling work (i.e. creating the input/output representations; building and testing the model) and all analyses contained within the thesis. For all chapters of the thesis, I produced initial drafts which were then developed further with guidance from my supervisors to produce the final versions.

CHAPTER 2

SEGMENTAL PHONOLOGY HAS ITS ORIGINS EXPLICITLY IN LETTERS BUT IMPLICITLY IN WORDS

Stephanie AINSWORTH

Stephen WELBOURNE

Anna WOOLLAMS

Anne HESKETH

Paper prepared for submission to Child Development

Abstract

Current theories of phonological development make contrasting predictions about the role of vocabulary growth and orthographic knowledge on the emergence of segmental phonological representation (PR). Testing these predictions in children is made difficult by the metacognitive nature of the tasks commonly used to assess phonological representations. The current study uses both implicit and explicit PR measures alongside measures of vocabulary and letter-sound knowledge (N=88, age range: 3;2 to 5;7). Results show that while explicit knowledge of phonological representations is related to letter-sound knowledge, implicit PR tasks provide evidence of segmental phonology which is related to vocabulary growth and is not mediated by orthography. This study thus shows the importance of tapping into implicit phonological representation using viable and practical tasks.

Introduction

Phonological representations are the way that children store the sound structure of words that they know. It is important to understand the nature and development of phonological representations as they are linked to children's explicit knowledge of phonological elements – known as phonological awareness or PA – which in turn has been linked to later reading success (e.g. Lonigan, Burgess & Anthony, 2000). Given the importance of raising literacy levels, an understanding of the processes leading up to the mastery of reading and how these break down in children with reading difficulties remains a priority (Hulme & Snowling, 2009; Department of Education, 2011). The extent to which the link between children's phonological knowledge and later reading success may be mediated by knowledge of letters has also been debated (Castles & Coltheart, 2004). Any comprehensive model of reading acquisition therefore needs to include a detailed account of how phonological representations develop and how they interact with children's explicit phonological awareness and their knowledge of letter-sound correspondences. Yet there is still little agreement as to the structure of children's phonological representations and the way that they evolve over time.

The study of children's phonological representations is complicated by the fact that the phonological awareness tasks often used to measure them tend to require a metacognitive awareness of sound structure. Measurements of children's explicit knowledge of their phonological representations (usually referred to as phonological awareness or PA tasks) can be contrasted with implicit PR measurements (such as similarity judgements) which do not require children to have any conscious knowledge of the sound segments within words. This study is the first to directly contrast these two types of task in order to test the different predictions offered by current models of phonological development.

Key accounts of PR development

The Accessibility Account

Some researchers have argued that young children's phonological representations are stored at a level of phonological detail similar to that of adults (e.g. Bailey & Plunkett, 2002, Ballem & Plunkett, 2005, Swingley & Aslin, 2000; Swingley, 2009) but that they can only consciously access the sound structure of words (e.g. in explicit phonological awareness tasks) once they have developed the required metacognitive skills and letter knowledge needed to tap into them (Lieberman & Shankweiler & Liberman, 1989; Rozin & Gleitman, 1977). Support for this account comes from studies where toddlers younger than two years have shown sensitivity to mispronunciations differing by only one phoneme

in preferential looking (Bailey & Plunkett, 2002; Ballem & Plunkett, 2005) and visual fixation tasks (Swingley & Aslin, 2000; Swingley, 2009). It is argued that as children are sensitive to very small changes in pronunciation then their phonological representations must be stored at a fine level of detail. According to this view then, implicit behavioural tasks that circumvent the need for metacognitive awareness of sound structure should show fine grained structure in children's phonological representations.

The Emergent Account

The second viewpoint – the emergent account – suggests that words are not initially stored in terms of phonemes and that adult-like phonological representations only emerge after a period of gradual lexical restructuring. Evidence for the emergent view comes from studies which indicate qualitative differences in the way that children and adults classify words (see Metsala & Walley, 1998 for a review) and studies which show representations to become more segmentally detailed over development (Metsala, 1997; Storkel, 2002; Carroll & Myers, 2011, Ainsworth, Welbourne & Hesketh, submitted). For example, Treiman and Breaux (1982) used a common phoneme task to show that while adults tend to classify words in terms of shared phonemes, children tend to make classifications based on global similarity (how much the words sound like each other overall). More recently Carroll & Myers (2011) found that while children and adults were sensitive to both shared manner of articulation and shared phonemic segments within classification tasks, the proportion of phonemic classifications increased over development.

Proponents of the emergent view argue that, although infants have been shown to be sensitive to small phonetic differences, this does not require their lexicon to be phonemically organised (Bowey & Hirakis, 2006; Ziegler & Goswami, 2005). Ziegler and Goswami (2005) point out that the ability to distinguish between two phones is not the same as recognising that different surface realizations of a given sound can be categorised as one phoneme (e.g. the /k/ in key and /k/ in 'car' are exemplars of the same phoneme). In other words, while early infant discrimination studies provide evidence that children store different phonemic sounds distinctly (i.e. the /t/ in ten is stored in a way that makes it separable from the /b/ in ben), they do not speak to the question of whether incidences of the same phoneme occurring within different words are stored similarly (e.g. whether /b/ is stored within 'big' in the same way as /b/ within 'ben').

Under the umbrella of the emergent view, three key variants have been proposed:

i) Vocabulary growth drives PR segmentation which is a prerequisite for PA

The lexical restructuring model, LRM (Metsala & Walley, 1998) suggests that vocabulary growth gradually stimulates the segmentation of phonological representations into onset-rime and then phoneme level representations. Within this account segmental representations are a prerequisite for explicit PA – in other words children’s representations need to be stored in terms of rime segments before they can access rime segments during explicit rime level PA tasks. Similarly, children’s representations need to be further segmented in terms of individual phonemes before they can access phonemic segments during explicit tasks measuring phoneme level PA. Evidence in support of this account comes from studies which show a developmental shift in sensitivity from the global properties of phonological word forms to their subcomponents (Metsala & Walley, 1998; Metsala, 1997). Metsala & Walley (1998) argue that this qualitative change in the way that children perceive words is indicative of phonological representations being stored initially in terms of global properties, then becoming increasingly componential as they learn more words.

Further evidence for lexical restructuring comes from studies which find that the phonological representations of words from dense phonological neighbourhoods (words with lots of neighbours within the lexicon that are phonemically similar, e.g. neighbours of cat might include can, cot, hat) are more segmental than for words from sparse neighbourhoods (words with few phonemically similar neighbours) (Metsala, 1997, Garlock, Walley & Metsala, 2001). This relation between neighbourhood density and PR segmentedness (the extent to which phonological word forms are stored at a particular grain size, e.g. in terms of onset and rime or individual phonemes) is explained within the LRM in terms of words from more dense neighbourhoods needing to be restructured first in order to keep them distinct from the large number of similar sounding words within their neighbourhood (Metsala, 1997; Metsala & Walley, 1998).

ii) Children’s phonological representations become phonemic only once they have learnt the mappings between graphemes and phonemes

Psycholinguistic grain size theory (Ziegler & Goswami, 2005) shares the idea with the LRM that phonological representations are gradually restructured to allow more efficient representation of similar sounding words as children’s vocabularies grow, but within this account restructuring is not framed in terms of a shift from large to small representational components. Rather it is proposed that detail is added to children’s

representations at all grain sizes as neighbourhoods of similar sounding words emerge within the lexicon. Another key difference between the two accounts is that while the LRM proposes that phonemes emerge naturally through spoken language experience, in grain size theory phonemic representation only emerges when children learn about phonemes explicitly, usually through the teaching of grapheme-phoneme correspondences. Evidence in support of this idea comes from studies which show that children lack the ability to count phonemes before they receive literacy instruction but develop the skill soon after literacy instruction begins (see Ziegler & Goswami, 2005 for a review).

iii) Lexical restructuring occurs in the absence of literacy, but letter knowledge is needed for explicit awareness of phonemes

Ventura and colleagues have investigated PR development by studying adults with varying levels of literacy (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007). They found that illiterate adults may develop fine grained segmental representations despite having limited letter knowledge, but they are not able to access them in the same way as literate adults on phoneme level PA tasks. Other studies within illiterate populations have shown similar difficulties with phoneme level PA tasks e.g. phoneme reversal ('what is 'pa' backwards?'), Adrian, Alegria & Morais, 1995) and phoneme deletion (e.g. 'when I say 'kur', you say 'ur', de Gelder, Vroomen & Bertelson, 1993), alongside normal performance on rime level PA tasks such as rhyme detection (e.g. 'do 'dak' and 'lak' rhyme?' de Gelder, Vroomen & Bertelson, 1993). Together these studies suggest that while phonemes may emerge through spoken language experience alone at a representational level, children need to learn about letters to be able to attend to the phonemic segments explicitly. Rime segments on the other hand appear to be accessible independent of letter knowledge.

Predictions made by the key theories of phonological development

The four theoretical accounts listed above – the accessibility account plus three variants of the emergent view – each make different claims about the relations between PR segmentedness, phonological awareness, vocabulary growth and letter-sound knowledge (used here to refer to the mappings between graphemes and phonemes generally not just single letter-sound mappings). The key predictions allowing us to distinguish between the four accounts are summarised in Table 2.1.

Table 2.1. Summary of predictions made by the four key theories of phonological development

Theoretical account	Early competence on implicit PR tasks	A unique relation between...			Letter-sound knowledge is prerequisite for...	
		PR segmentedness and vocabulary	PR segmentedness and letter-sound knowledge	PA and letter-sound knowledge	Phoneme level PRs	Phoneme level PA
Accessibility view ^a	Yes	No	No	Yes	No	Yes
Lexical Restructuring Model ^b	No	Yes	Unspecified	Unspecified	No	Unspecified
Psycholinguistic Grain Size Theory ^c	No	Yes	Yes	Yes	Yes	Yes
Restructuring but not phoneme awareness occurs in the absence of literacy ^d	No	Yes	No	Yes	No	Yes
Current study	No	Yes	No	Yes	No	Yes
Evidence from current study	Young nursery children are only just above chance on implicit PR measures.	Implicit PR measures are predicted by explicit PR and vocabulary but not age or letter-sound knowledge (Table 2.5)		Explicit PR is predicted by implicit PR and letter-sound knowledge but not age or vocabulary (Table 2.6)	Children with little letter-sound knowledge show sensitivity to phonemes on implicit PR tasks (Table 2.7)	Children with little letter-sound knowledge can blend phonemes but cannot explicitly tell you the sounds within a word (Table 2.7)

^a e.g. Rozin & Gleitman (1977), Liberman, Shankweiler & Liberman (1989), ^b Metsala & Walley (1998), ^c Ziegler & Goswami (2005), ^d Ventura, Kolinsky, Fernandes, Querido & Morais (2007)

The first row relates to predictions made by the accessibility account. Given that this viewpoint argues phonological representations to be fully segmental from infancy, we would expect early competency on tasks measuring phonological representation. We would also expect no unique relation between PR segmentedness and either vocabulary or letter-sound knowledge. Nor would we expect letter-sound knowledge to be needed for phoneme level representations. We would however expect a unique relation between letter-sound knowledge and PA and for letter-sound knowledge to be prerequisite for the emergence of phoneme level PA given that the accessibility view proposes that experience with an orthography unlocks explicit access to the stored sound segments within words (Rozin & Gleitman, 1977; Liberman, Shankweiler & Liberman, 1989).

All three emergent theories predict that as children learn more words the lexicon is forced to represent words componentially in order to keep similar sounding words distinct – thus we would expect PR segmentedness to be uniquely related to vocabulary in each case. We would also expect performance on implicit measures of PR to increase gradually in children rather than full competency early on. Where these variants differ is in the predictions they make about the role of letter-sound knowledge. The lexical restructuring model makes no specific predictions about the potential influence of orthographic knowledge on children's phonological representations or on their explicit awareness of them (PA), however it does suggest that phonological representations become phonemic through oral language experience alone and therefore letter-sound knowledge is presumably not necessary for phoneme level representations (Metsala & Walley, 1998).

Psycholinguistic grain size theory (Ziegler & Goswami, 2005) suggests that letter-sound knowledge precipitates phoneme level representation via the emergence of an explicit awareness of phonemes (i.e. children usually become conscious of phonemes when they are taught the links between letters and sounds). It therefore predicts a unique relation between letter-sound knowledge and both PR segmentedness and PA, as well as a need for letter-sound knowledge before children can succeed on measures of both PR segmentedness and PA tasks at the phoneme level.

Finally Ventura et al.'s account (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007) predicts that phonological representations can be restructured in the absence of letter knowledge but that letter knowledge is needed to gain access to the components within the phonological representations explicitly during PA tasks. Within this account we would therefore expect no unique relation between letter-sound knowledge and PR segmentedness but a significant unique relation between letter-sound knowledge and phonological awareness.

One way to test these predictions would be to take concurrent measures of each of these constructs (PR segmentedness, PA, vocabulary and letter-sound knowledge) within a sample of children over an age range where their letter-sound knowledge and explicit awareness of phonological segments is emerging. However, as highlighted earlier, it is not possible to measure PR segmentedness directly. Instead the current study uses novel implicit measures of phonological representation which do not require children to have any explicit knowledge of phonological segments (as most PA tasks do) and may therefore allow us to get closer to the phonological representations themselves than is possible when using traditional PA tasks. By contrasting implicit and explicit measures of phonological representation (usually called PA in the latter case) we are able to separate children's implicit sensitivity to phonological segments (a proxy for the underlying segmentedness of their phonological representations) from their conscious awareness of these segments, thus allowing us to test the predictions made by the key theories of PR development.

Similarity classification as an implicit measure of phonological representation

One way to probe children's phonological representations implicitly is to ask children to classify words in terms of how similar they are (Storkel, 2002, Carroll & Myers, 2011, Ainsworth, Welbourne & Hesketh, submitted). The rationale here is that if children classify words in terms of shared segments at a particular grain size (e.g. phonemes) we can infer that their lexicons are organised in terms of these segments. Similarity based classification tasks are distinct from phonological awareness tasks in that they do not require children to explicitly reflect on the phonological segments within words, rather they are asked to make an implicit 'sounds like' judgement.

Previous studies have used similarity judgement tasks to measure PR segmentedness alongside letter-sound knowledge (Carroll & Myers, 2011) and/or vocabulary size (Ainsworth, Welbourne & Hesketh, submitted; Storkel, 2002; Carroll & Myers, 2011) but have not included measures of explicit phonological awareness. In order to test the predictions made by the key accounts of phonological development, there is a need for studies which contrast implicit measures of phonological representation, which do not require an explicit knowledge of the phonological system, with explicit PR measures (often called PA) which measure children's overt awareness of sound segments. This study uses concurrent implicit and explicit PR measures alongside measures of vocabulary growth and letter-sound knowledge within a sample of nursery and reception children who we would expect to have a range of PA and letter-sound knowledge to test the predictions made by the key theories of phonological development (Table 2.1).

Method

Participants

90 children were recruited from two UK mainstream primary schools with low and medium sociodemographic status as indexed by the percentage of free school meals (based on the 2012 school dashboard data, see Ofsted, 2014). Children's ages ranged from 3;2 to 5;7, with 48 children in nursery and 42 children in reception class (in the UK most children enter school in the nursery class aged 3 to 4 years). The gender balance was 51 boys and 39 girls. To be included in the study children needed to have at least one English speaking parent, to have no known history of speech or hearing problems (as reported by the teacher) and to not be on the special educational needs register for any behavioural or developmental concerns. While the younger nursery (aged 3;2 to 3;10) and younger reception (aged 4;0 to 4;7) groups were both tested in the Autumn term, the older nursery (aged 4;0 to 4;5) and older reception (aged 4;7 to 5;7) groups were tested in the late Spring/early Summer terms.

An adult control group was also recruited to confirm the construct validity of the novel PR measures: we wanted to check that literate adults who presumably have detailed fine-grained representations do in fact score highly on the implicit measures of phonological representation. The adult sample consisted of 74 undergraduate students. Adult participants were excluded if English was not their first language or if they had any diagnosed dyslexia or hearing difficulties. The adult group was tested on the five implicit PR measures and the explicit rhyme task. They were not tested on blending, phoneme isolation and letter-sound knowledge tasks as it was assumed that they would perform at ceiling on these measures. They were also not tested on the vocabulary measures, which were suitable for children only.

Procedure

Children were tested in a quiet room within school, over 5 to 10 sessions depending on their age/attention span. All tasks requiring a verbal response from the child were audio-recorded. No corrective feedback was given except for one training item at the beginning of each task (with the exception of the expressive vocabulary and letter sound knowledge tests which did not have any training items). General praise and stickers were given as encouragement regardless of performance.

Materials

Pictures were chosen to be familiar to young children. Almost all of the lexical items (98%) used are present within Storkel and Hoover's database (Storkel & Hoover, 2010) which was drawn from corpora of kindergarten and first grade children. To make sure that children identified the pictures correctly they were asked to name the pictures at the beginning of each trial and were told the correct name where necessary.

Within all the multiple choice implicit and explicit PR tasks the distracters were matched listwise in terms of frequency and two measures of phonotactic probability (positional segment average – which measures how often the segments within a word occur in that position within other words and biphone average which measures the frequency of pairs of sound segments) using Storkel and Hoover's online calculator (Storkel & Hoover, 2010). The full stimuli list including matching characteristics is provided in Appendices 1 to 7.

Implicit measures of phonological representation

Five implicit PR measures were devised, four of which probe segmentedness at both the onset-rime and phoneme level and one which probes PR segmentedness at the onset-rime level only. A summary of all tasks is given in Table 2.2.

Similarity based classification tasks

The first three tasks (listed within Table 2.2) involve children making similarity classifications and share the same rationale. Children are asked to compare auditory CVC stimuli in terms of how similar they sound. For each trial the target response is a word/pseudoword which shares two phonemes with the stimulus (e.g. *tet*-ten) and the closest distracter is a word/pseudoword which only shares one phoneme with the stimulus but is matched in terms of global similarity (e.g. *tet*-tape, see below for further details on global similarity matching). The idea here is that if children choose the closest segmental match more often than the globally matched distracter, then we can infer that they are sensitive to the number of shared segments over and above how similar the words/pseudowords are in terms of overall 'sounds-likeness'. Sensitivity to shared segments in turn suggests phonological representations are segmented at least at the grain size of those segments. In other words if children are sensitive to shared rimes then their representations must be segmented at least at the onset-rime level; if children are sensitive to shared phonemes within words this suggests a more fine grained phonemic representation of those words. The three tasks contain two types of item designed to tap into PR segmentedness at both the rime and phoneme levels. While rime level items

involve two shared phonemes in the rime position (e.g. *tain-rain*), phoneme level items have two shared phonemes in body position (e.g. *tet-ten*). It is assumed that while rime items can be completed with representations which are segmented at the rime level only (by comparing the rime segments as a whole), a more fine grained phonemic representation is needed for phoneme level items where the rime segment is disrupted (see Metsala & Walley, 1998 and Ziegler & Goswami, 2005 for discussion of the developmental primacy of the rime in English).

Global similarity matching

Within the similarity based measures the closest distracter (or only distracter in the case of the two-choice tasks) was matched to the segmental (target) response in terms of global similarity. For example when asking children whether '*nig*' or '*teg*' sounds the most like 'pig', although '*nig*' is segmentally closer to 'pig' (it shares two rather than one phoneme with 'pig') '*teg*' is equally close to 'pig' in terms of global similarity or overall 'sounds-likeness'. Global similarity was operationalised using adult ratings collected by Singh and colleagues (Singh & Woods, 1971; Singh, Woods & Becker, 1972). Scores of how dissimilar the standard was from the target and the distracters were calculated using the same concatenative method adopted by Treiman and Breaux (1982) and Carroll and Snowling (2001). For example the dissimilarity score between the words 'pin' (pronounced /pɪn/) and 'bed' (pronounced /bed/) is the dissimilarity of /p/ and /b/ (3.9) plus the dissimilarity of /ɪ/ and /ɛ/ (2.22) plus the dissimilarity of /n/ and /d/ (4.8). The five measures of implicit phonological representation are illustrated in Figure 2.1.

Table 2.2. Summary of measures. All PR measures with the exception of phoneme isolation are receptive (involving picture selection rather than a spoke response)

Type of measure	Name of measure	Task description	Number of items		Chance level (%)
			Rime level	Phoneme level	
Implicit PR measures	Mispronunciation reconstruction	Puppet says a mispronounced word, (e.g. 'tel'). 'Which picture do you think the puppet is trying to say?'	6	6	25
	Pseudoword similarity	'Sen is going in the spaceship. Who sounds the most like Sen – is it Ses or Sif?'	8	8	50
	Mispronunciation conflict	Two monsters mispronounce a word (picture on screen, e.g. 'net'). 'Which monster said it the best? Which one sounded the most like it?'	8	8	50
	Incomplete word	Puppet says a word onset e.g. 'cl'. Child guesses what he wants from that sound,	8	8	25
	Implicit rhyme	Puppet says three rhyming words (e.g. 'Fun, done, run...'). 'What comes next? Which one finishes the pattern?'	8	0	25
Explicit PR measures	Explicit rhyme	'Which one rhymes with...?' (e.g. run)	8	0	25
	Blending	Robot says 't-e-n'. Child points to the corresponding picture.	8	8	25
	Phoneme isolation	Child is asked to say the sounds in a word, (e.g. 'cat' is 'c-a-t')	8	8	Free choice
Background language measures	Letter-sound knowledge	Child is shown a grapheme and asked what sound it makes.	35 items: 26 single graphemes, 8 digraphs and 1 trigraph		Free choice
	Word finding	Standardised test of expressive vocabulary.	no. of items is dependent on performance		N/A
	BPVS	standardised test of receptive vocabulary			

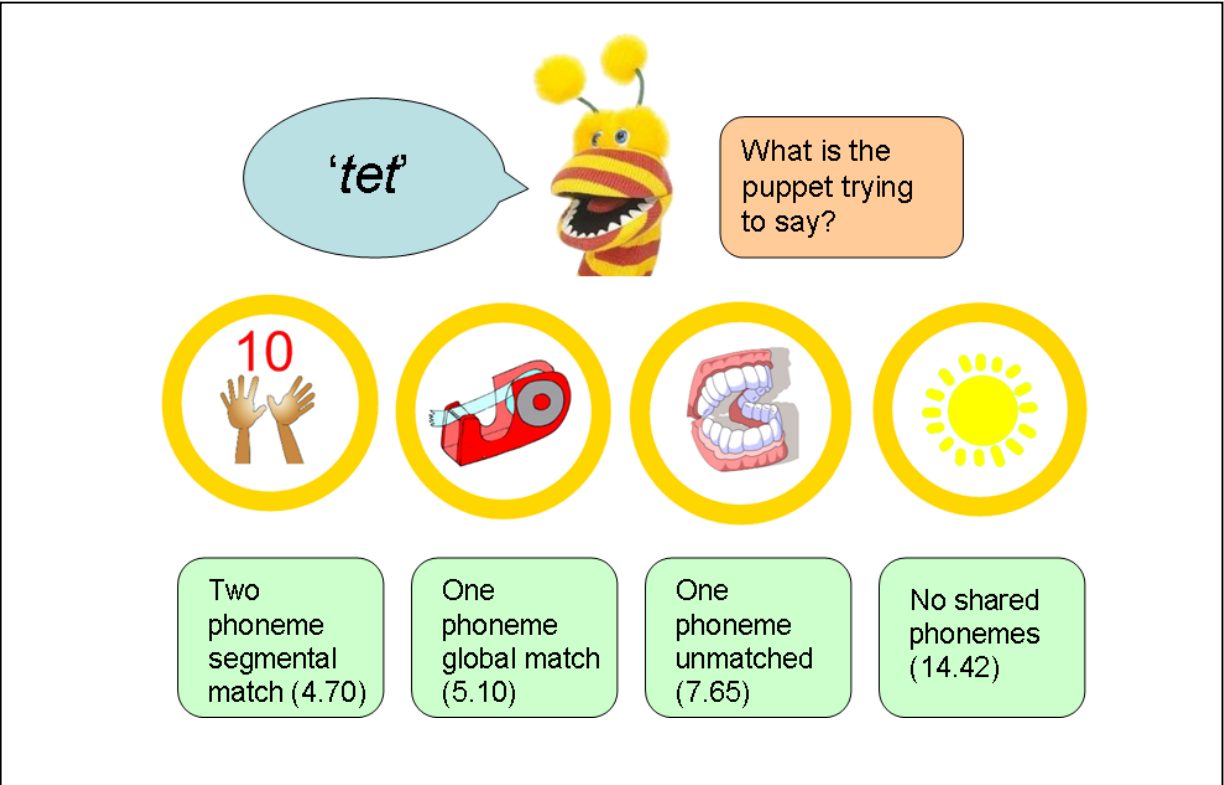


Figure 2.1a. Example phoneme level item from the mispronunciation reconstruction task.

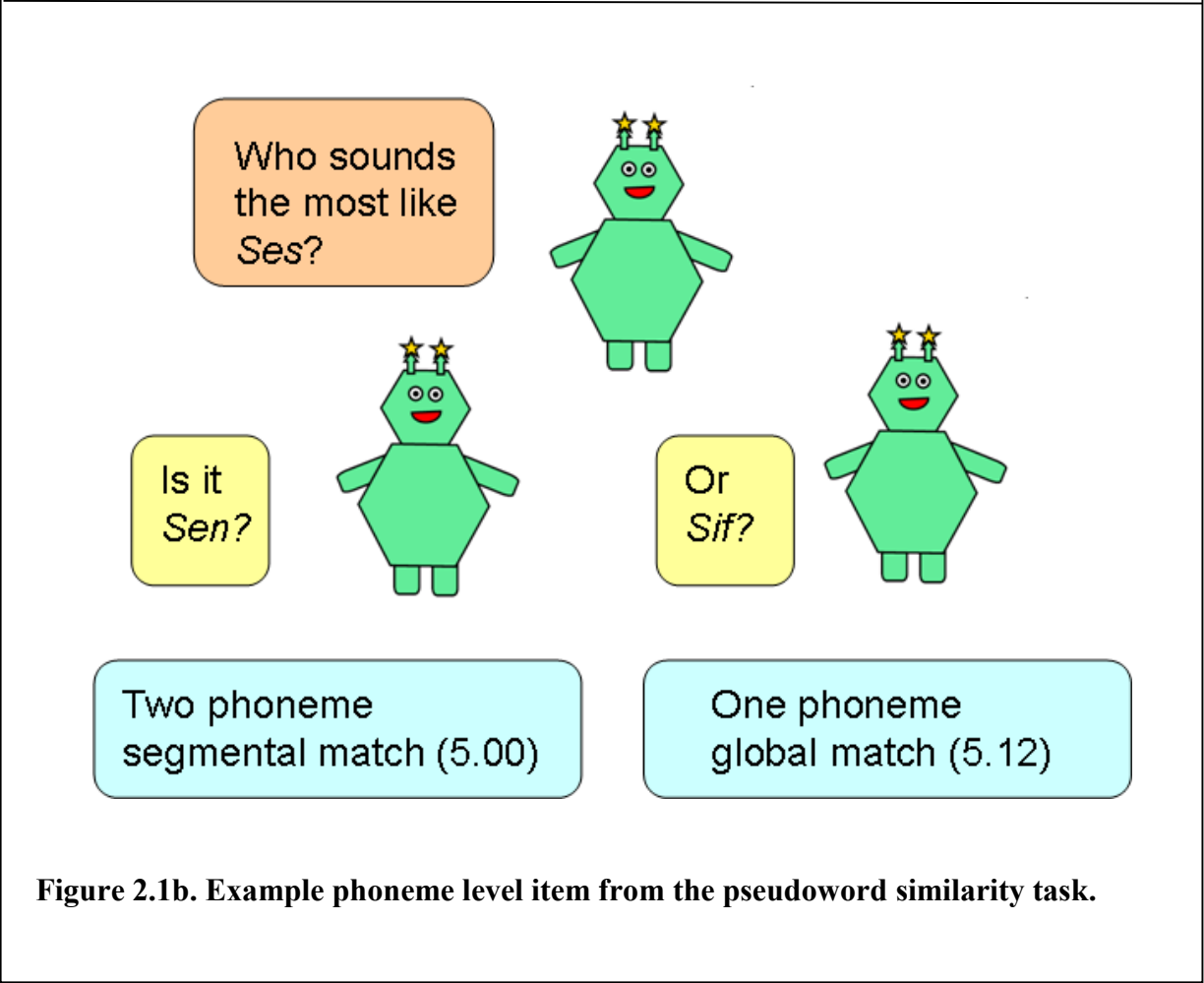


Figure 2.1b. Example phoneme level item from the pseudoword similarity task.

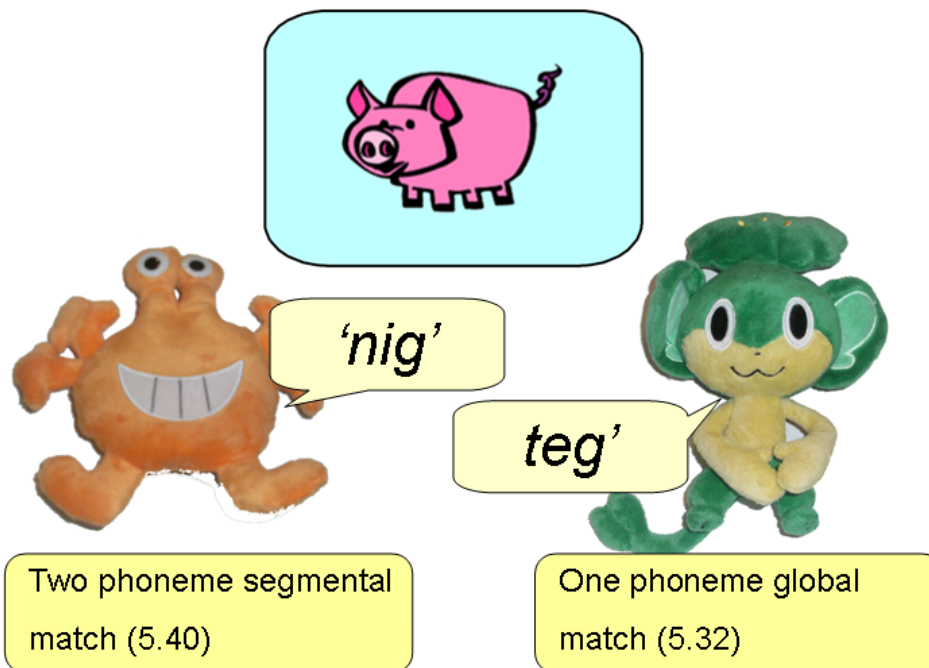


Figure 2.1c. Example rime level item from the pseudoword similarity task.

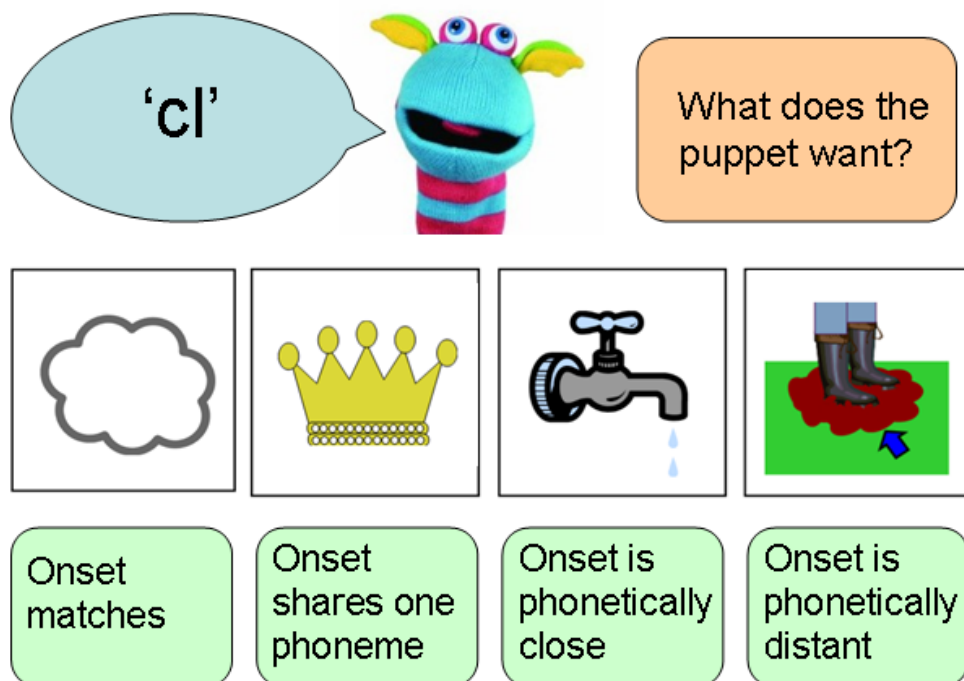


Figure 2.1d. Example phoneme level item from the incomplete word task.

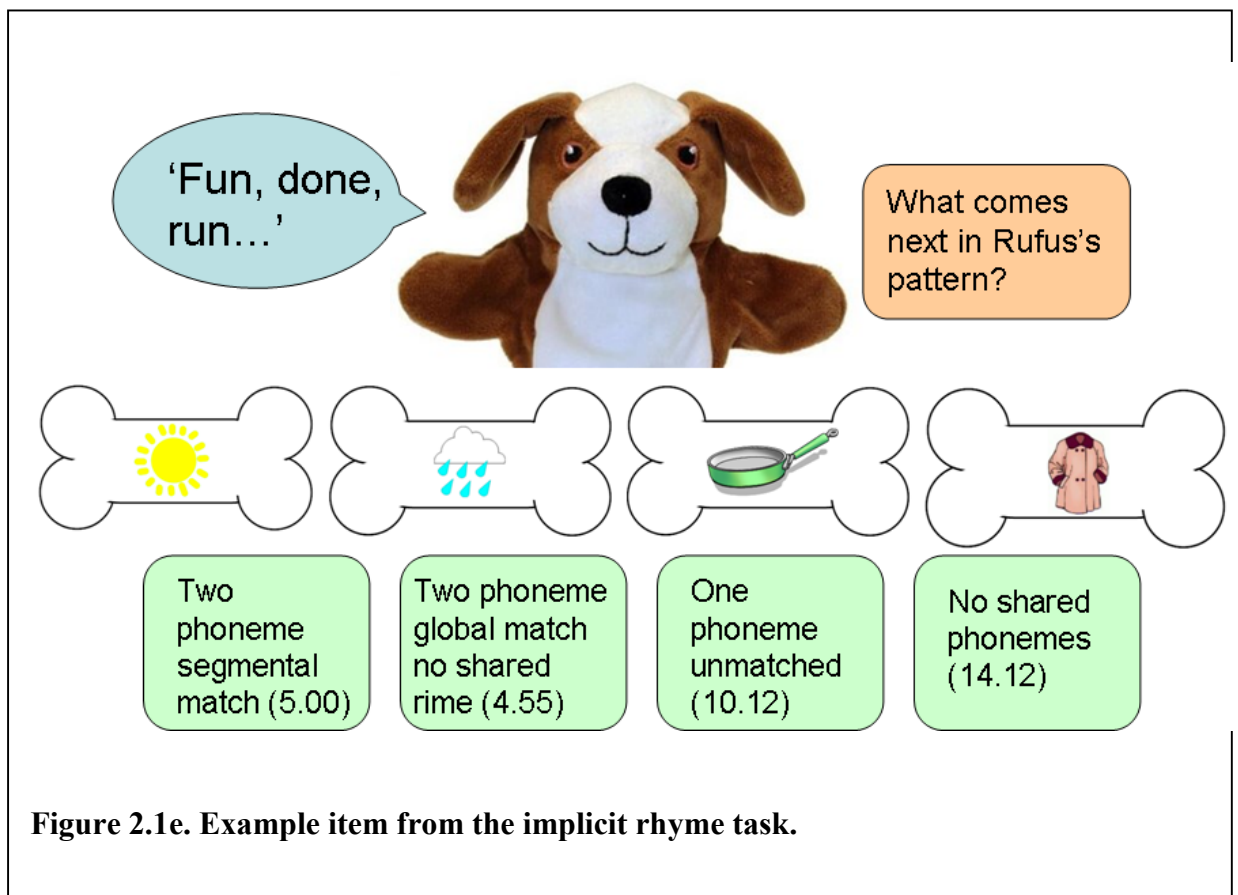


Figure 2.1e. Example item from the implicit rhyme task.

Figure 2.1. Implicit measures of phonological representation. The numbers in brackets show the global similarity distance between the stimulus and each response choice.

Mispronunciation reconstruction Children heard a puppet mispronounce a word and were then asked to guess which picture he was trying to say – which picture did it sound the most like (see Figure 2.1a) For example the puppet says ‘*tet*’ and the children choose whether he was trying to say ten, tape, teeth or sun.

Pseudoword similarity Children were asked which of two pseudoword alien names sounded the most like a third pseudoword name (see Figure 2.1b). For example, ‘Which one sounds the most like *Ses*, is it *Sen* or *Sif*?’ This task is an adaptation of the common phoneme classification task developed by Treiman and colleagues (Treiman & Breaux, 1982; Treiman & Baron, 1981). Although the current task has a similar rationale to the original task, the working memory demands are reduced by asking children to choose one name rather than the most similar pair of names. It also takes a more continuous approach to phonemic similarity (with the two choices sharing one and two phonemes rather than one and no phonemes) to make it more comparable with the other tasks within the battery.

Mispronunciation conflict In this novel task children heard three aliens attempt to say a word (with the target picture shown on screen) and were first asked to choose which one said the word correctly. This part of the task was included to check that the children were able to recognise the correct form of the word as distinct from the mispronunciations. They then listened to the two aliens who said the word incorrectly again and were asked, ‘Which one said it the best? Which one sounded the most like it?’ (see Figures 2.1c) For example children chose whether ‘*nig*’ or ‘*teg*’ sounds most like ‘*pig*’.

There are two further implicit PR tasks:

Incomplete word Children heard an onset and were asked which picture (out of a choice of four) the puppet wanted, where the target picture was of a word sharing the onset spoken by the puppet (see Figure 2.1d). For rime level items, the stimulus was a single consonant (e.g. /b/ with target picture ‘boat’). For phoneme level items children were required to link a consonant cluster to a word (e.g. /cl/ - ‘cloud’), where the closest distracter began with a cluster sharing the same initial consonant (e.g. ‘crown’). The rationale here is that while children can link /b/ to ‘boat’ when their phonological representation for ‘boat’ is segmented at the onset rime level only, children will only be able to distinguish between the /cl/ in ‘cloud’ and the /cr/ in ‘crown’ if they have a finer grained phonemic representation of these words.

Although this task is superficially similar to the phoneme identification task (Elbro & Petersen, 2004) which measures children’s explicit phonological awareness (PA), there is a key difference which makes it a measure of PR segmentation rather than PA. Unlike PA tasks the incomplete word task does not ask children to explicitly reflect on the sounds in words, rather they are asked to simply guess what the puppet wants, having heard it say a sound. In this way the incomplete word task is designed to be as implicit as possible, so that we can tap into the segmentedness of children’s phonological representations without requiring children to have any explicit knowledge of rimes or phonemes.

Implicit rhyme Children heard a puppet say three rhyming words (e.g. fun, done, run) and were then asked to choose which picture completes the pattern (e.g. sun, see Figure 2.1e). Unlike the explicit rhyme task described below, the implicit rhyme task purposely avoids the use of the word ‘rhyme’ and does not ask children to explicitly reflect on the shared rime segments. Rather children are simply required to pick up on the rhyme pattern implicitly.

Explicit measures of phonological representation

Explicit rhyme Children were shown four pictures (e.g. sun, rain, pan, coat) and asked which one rhymes with a given stimulus (e.g. ‘Which one rhymes with run?’). The same stimuli, target items and distracters used in the implicit rhyme task were also used here to allow a direct comparison across implicit and explicit measures of phonological representation at the same level of segmentedness (i.e. the onset-rime level). The explicit rhyme task always took place after the implicit rhyme task with a gap of a few days in between.

Blending Children heard a robot (voice was pre-recorded) say either an onset and a rime (e.g. t-en) or three individual phonemes (e.g. t-e-n) and were asked to select the corresponding picture.

Phoneme isolation Children were shown a picture and asked to say the sounds in the word (e.g. c-a-t if shown a picture of a cat). To avoid unnecessary testing, if children were not able to isolate any of the sounds in eight consecutive words, testing was stopped at this point and it was assumed that children would have scored zero on all remaining items (this criterion applied to 19 children).

Background language measures

Letter-sound knowledge Children were shown a grapheme (a letter or group of letters) and were asked what sound it makes. The graphemes were presented in order of difficulty as indexed by the order in which the Letters and Sounds framework (PNS, 2007) recommends that they are taught. Testing was halted if children failed to say the sounds for 8 consecutive letters and the child was scored as knowing the number of letters answered correctly up to this point (this criterion applied to 44 children). It should be noted that within this paper when we use the term ‘letter-sound knowledge’ we refer broadly to refer to all grapheme-phoneme correspondences (e.g. including the mappings for digraphs like ‘sh’ – /ʃ/), not just to the mappings between single letters and sounds.

Vocabulary Expressive and receptive vocabulary were measured using the Renfrew Word Finding Vocabulary Test (Renfrew, 1995) and British Picture Vocabulary Scale, BPVS (Dunn, Dunn, Sewell, Styles, Bryzka, Shamsan & Burge, 2009) respectively. Standardised procedure for the BPVS was adapted slightly: all children started at set 1 regardless of age so that raw scores could be compared between age groups.

Counterbalancing

For all of the implicit and explicit PR measures, two sets of item orders were used to control for order effects. Within each age group children were randomly assigned to either set. The order in which the tasks were delivered was kept constant across all participants. This was done to ensure that the explicit tasks always followed the implicit tasks. If children carried out the explicit tasks first this may have impacted on their implicit performance due to increased salience of the segments accessed within the explicit tasks. Given that we were interested in separating children's implicit and explicit knowledge as much as possible, it therefore made sense to conduct the explicit measures after all the implicit measures had been completed.

Results

Data screening

Data screening showed one adult participant to have extreme outliers for three of the six tasks (scores with values more than 1.5 times the interquartile range below the first quartile). This participant's scores were removed prior to analysis. No extreme outliers were found within the child data, but two children did not attempt one or more of the tasks. The scores from these two children were also excluded listwise from the analyses that follow.

Summary of children's performance

The mean scores for the implicit and explicit PR tasks in Figures 2.2 and 2.3 show children to be significantly above chance on all multiple choice tasks from the older nursery group upwards. This indicates that children aged 4;0 to 4;5 in the second half of their nursery year are already showing implicit sensitivity to shared segments in words and have some explicit phonological awareness. For the phoneme isolation task, which was a free choice task a success threshold of 42% was chosen following an inspection of the histogram for phoneme isolation performance. Children did not succeed on this task until reception and they only achieved high levels of performance in the second half of the reception year. Nursery children aged 3;2 to 3;10 tested at the beginning of the school year had limited explicit phonological awareness (only scoring above chance on the blending task) and knew very few letter-sound correspondences (mean number of letter sounds known=3, SD=6), yet scored just above chance on 3/5 of the implicit PR tasks. This suggests that some sensitivity to shared segments may develop before children learn about letters and become explicitly aware of phonological segments. Implicit PR performance of the youngest children was however only just above chance and remained below ceiling even at the end of reception.

Performance on the implicit tasks where children were given four choices on each trial (Figure 2.2a) shows children's implicit segmental sensitivity to increase steadily over the developmental period studied. Performance on the two-choice implicit PR tasks (Figure 2.2b) remains relatively flat with a significant between groups difference found only for the pseudoword similarity task between the young and old nursery children ($t(44)=2.87$, $p=.006$). This apparent plateau in segmental sensitivity on the mispronunciation conflict and pseudoword similarity tasks may be due to the reduced sensitivity inherent in tasks with a 50% chance level relative to a 25% chance level. Adult performance across all implicit measures was high confirming that the implicit tasks do elicit a segmental response in literate adults who presumably have fine grained phonological representations.

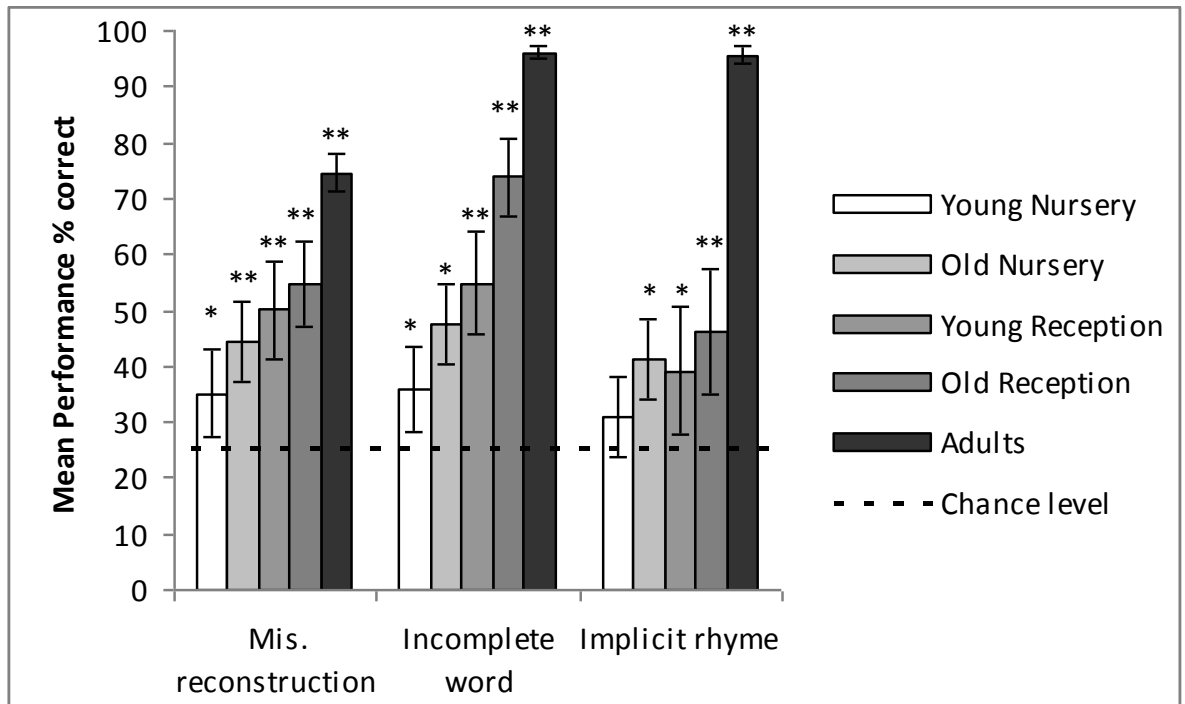
Performance of the children at the end of reception remained significantly below adult levels on all tasks (t values ranged between 4.80 and 9.81, $p < .001$).

How are PR segmentedness, phonological awareness, vocabulary and letter-sound knowledge related to one another?

In order to investigate the relation between PR segmentedness and the other key variables implicated in theories of phonological development we first reduced the battery of scores into 3 weighted averages plus the letter-sound knowledge score which remained as a stand-alone variable. The first weighted average or component was generated by entering the five implicit PR measures into a principal components analysis. The second and third components were generated in the same way from the three explicit PR scores and the two vocabulary measures respectively. In this way we were able to condense the data into four key variables: implicit PR (children's implicit sensitivity to the segmental structure of their phonological representations which we are using as a proxy for PR segmentedness), explicit PR (children's explicit knowledge of the segmental structure of their phonological representations), vocabulary and letter-sound knowledge (see Table 2.3).

All four variables increased over the developmental period sampled, as shown in Figure 2.4. The fact that all skills are developing concurrently poses a challenge in terms of evaluating exactly how the four constructs are related. As a first step a correlational analysis was carried out with age partialled out to control for general developmental factors such as attention span. All correlations (Table 2.4) were found to be highly significant confirming that these 4 measures of theoretical interest are empirically related independent of age. To further investigate the nature of these interrelations and test the predictions made by the different theories of PR development (see the columns two to four of Table 2.1) two regression models were created.

a)



b)

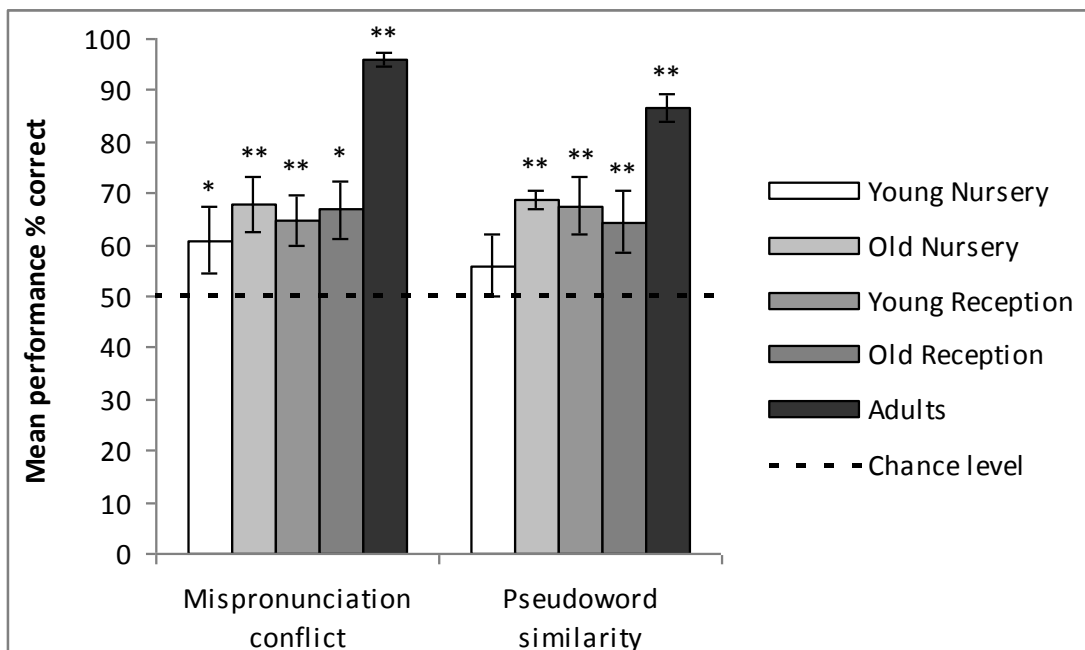


Figure 2.2. Performance on implicit PR measures with a chance level of a) 25% and b) 50%. Asterisks indicate performance significantly above chance at the $p < .05$ (*) and $p < .001$ () level and error bars represent confidence intervals of 95%.**

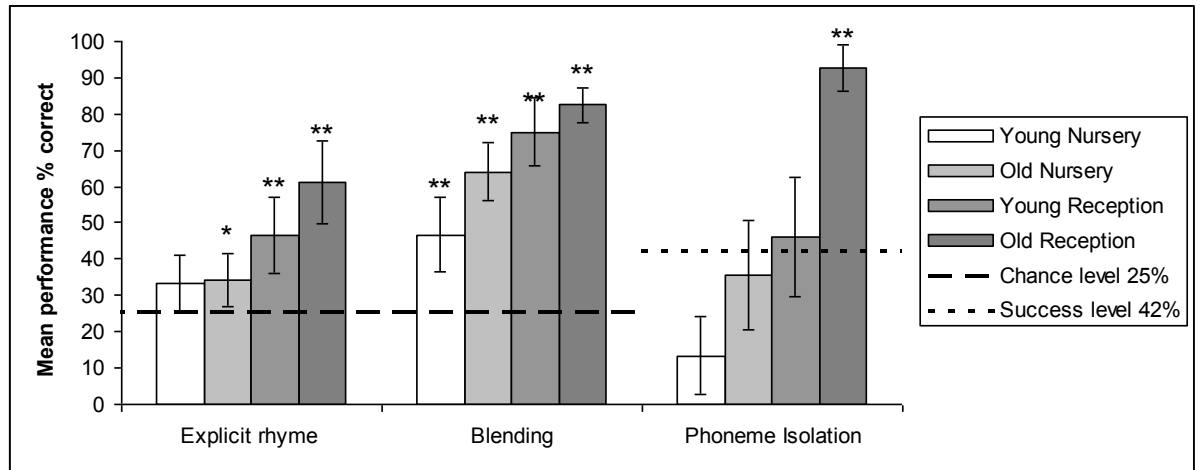


Figure 2.3. Performance on explicit measures of phonological representation.

Asterisks indicate performance significantly above chance at the $p < .05$ (*) and $p < .001$ (**) level and error bars represent confidence intervals of 95%.

Table 2.3. Loadings for the implicit PR, explicit PR and vocabulary PCA components

Implicit PR		Explicit PR		Vocabulary	
Task	Loading	Task	Loading	Task	Loading
Mis. reconstruction	.79	Explicit rhyme	.70	Word finding	.94
PW similarity	.70	Blending	.87	BPVS	.94
Mis. conflict	.69	Phoneme isolation	.91		
Incomplete word	.79				
Implicit rhyme	.61				
$R^2 = .52$		$R^2 = .69$		$R^2 = .87$	

Within the first regression model implicit PR was entered as the outcome variable with age, explicit PR, vocabulary and letter-sound knowledge entered simultaneously as the predictors. Within the second regression model explicit PR was the outcome variable with the other four variables: age, implicit PR, vocabulary and letter-sound knowledge entered simultaneously as predictors. The decision to use multiple regression in this way was made to allow us to investigate the relative influence of vocabulary and letter-sound knowledge on measures of implicit and explicit phonological representations. Implicit PR was included as a predictor of explicit PR and vice versa, given that we would expect children's implicit and explicit knowledge to be related to one another. On the one hand, the lexical restructuring model predicts that the level of segmentation of children's

phonological representations sets the stage for children's explicit rime and phoneme awareness, but we might also expect explicit knowledge of phonological segments to boost performance on the implicit measures due to increased salience of the segmental response choice. Age was also included as a predictor within both models, in an attempt to isolate the influence of vocabulary and letter-sound knowledge on PR performance independent of general age dependent factors (e.g. attention span, ability to follow instructions, etc.). The two regression models were also conducted using a stepwise method, with age entered within the first step (rather than simultaneously). The same pattern of results was obtained (as reported in Tables 2.5 and 2.6).

The authors acknowledge that the use of two linear regression models potentially oversimplifies the relationships between the variables of interest. Path analysis, where the interrelations between all four variables could have been investigated, would have provided a more sophisticated tool for analysing the data; however a much larger sample size would have been required to yield meaningful results from this type of analysis.

Within the first model, vocabulary and explicit PR were found to be significant predictors of implicit PR, whereas age and letter-sound knowledge were not (see Table 2.5). The model yielded an adjusted R^2 of 0.48. Within the second model age and vocabulary were not significant predictors of explicit PR, whereas implicit PR and letter-sound knowledge were (see Table 2.6). This model explained a greater proportion of the variance with an adjusted R^2 of .74. The regression models presented above have allowed us to test the predictions made within the columns two to four of Table 2.1. These results support Ventura's proposition (Ventura et al., 2007) that letter-sound knowledge may not be needed for segmentation of the underlying representations themselves, but is important for the development of explicit phonological awareness.

As noted above, the use of two separate regression models is limited in that it represents an oversimplification of the relationships between the variables and does not allow us to fully isolate the individual relationships between the four key variables. For example, it is possible that the strong relationship between explicit PR and implicit PR is diluting the apparently non-significant effect of letter-sound knowledge on implicit PR performance. It is encouraging however that the pattern of results is consistent with both the longitudinal data reported within Chapter 3 of this thesis and with the results of Ventura and colleagues' adult study (Ventura et al., 2007)

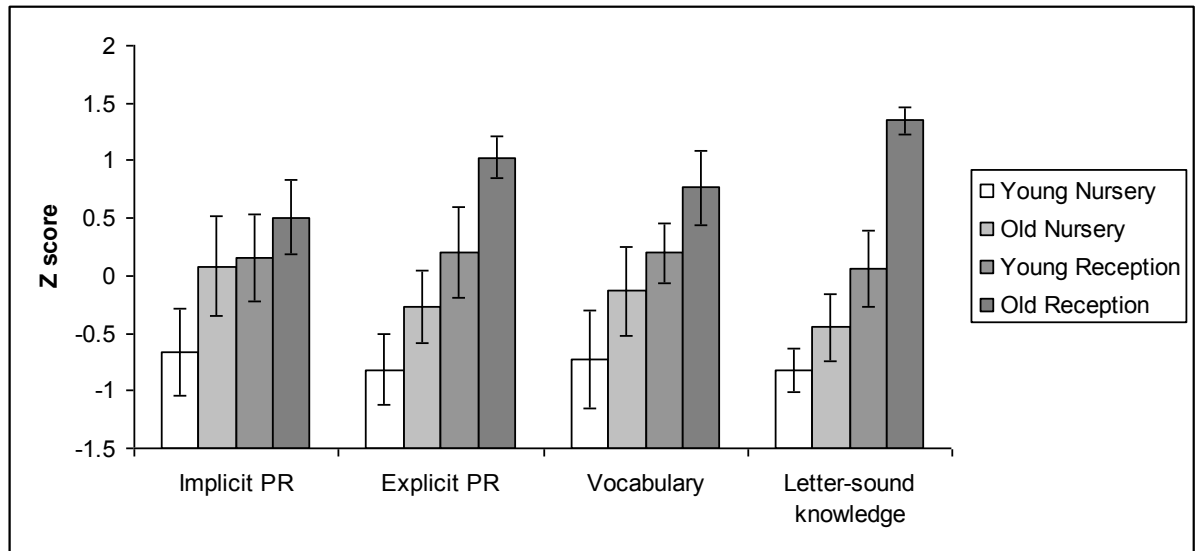


Figure 2.4. Clustered bar chart showing increasing performance on all 4 measures with age with error bars representing confidence intervals of 95%. The z scores for implicit PR, explicit PR and vocabulary are the factor scores arising from the PCA analysis. The letter-sound knowledge z scores were converted from the raw scores against a normal distribution.

Table 2.4. Partial correlations between implicit PR, explicit PR, vocabulary and letter-sound knowledge when controlling for age ($p < .0005$ for all correlations)

	Implicit PR	Explicit PR	Vocabulary
Explicit PR	.59	-	
Vocabulary	.47	.48	-
Letter-sound knowledge	.40	.65	.45

Table 2.5. Regression model with implicit PR as the outcome variable and age, explicit PR, vocabulary and letter-sound knowledge as the predictors

	B	SE B	Beta	Part correlations
Constant	.56	.84		
Explicit PR	.59	.14	.59**	.32
Vocabulary	.28	.11	.28*	.20
Letter-sound knowledge	-.004	.01	-.05	-.03
Age	-.01	.02	-.07	-.04

* $p < .05$, ** $p < .0005$ $R^2 = .50$

Table 2.6. Regression model with explicit PR as the outcome and age, implicit PR, vocabulary and letter-sound knowledge as the predictors

	B	SE B	Beta	Part correlations
Constant	-1.08	.59		
Implicit PR	.30	.07	.30**	.23
Vocabulary	.09	.08	.09	.07
Letter-sound knowledge	.04	.01	.54**	.31
Age	.01	.01	.07	.05

*p<.05, **p<.0005 R²=0.75

Do letter-sound mappings need to be learnt before phoneme level phonological representations emerge?

The final column of predictions relates to whether or not letter-sound knowledge is needed for phonological representations to become phonemic. In order to test this hypothesis children were divided into two groups with a threshold of less than three letters known assigned to the low letter knowledge group. Children were also grouped in terms of their performance on each of the implicit PR measures. Three different classifications were made corresponding to: performance overall on the task, performance on rime level items only and performance on phoneme level items only. The rationale for examining performance on items measuring implicit PR sensitivity at the rime vs phoneme level separately (rather than looking at overall implicit PR scores only) is that the theoretical accounts (Table 2.1) differ in terms of whether letter-sound knowledge is needed to reach the most fine grained level of representation, i.e. the phoneme.

In each case the classification threshold represents the chance level. The percentage of children within the low letter knowledge group who were above chance on each implicit PR measure was then calculated (see Table 2.7). We can see from Table 2.7 that the majority of children who knew fewer than three letters were above chance overall on 4 out of 5 of the implicit tasks. The implicit PR task which children in the low letter knowledge group found the hardest was the implicit rhyme task, with less than half the children found to perform above chance. For the other four tasks the percentage of low letter knowledge children who were above chance ranged from 53% to 69% (based on overall scores). This provides further support for the idea that while letter-sound knowledge is correlated with PR segmentedness (as evidenced in Table 2.4) it is not a prerequisite for segmentation to occur. The fact that the percentages of children in the low letter group who are above chance remains relatively high (36 % to 72%) when we consider phoneme level items

only, rules out the possibility that the sensitivity on implicit PR tasks seen within the low letter knowledge group is driven mainly by good performance on the rime level items.

Do we need letter-sound knowledge for explicit phonological awareness?

The regression models presented above suggest that while letter-sound knowledge may not be needed for children's underlying phonological representations to become segmented, it may play a key role in the development of explicit phonological awareness. If this is the case we would expect to find few or no children within the low letter group to be successful on explicit measures of phonological representation. To test this prediction, children were classified according to whether or not they succeeded on each of the three explicit PR tasks (Table 2.7). For the two multiple choice tasks the threshold corresponded to the chance level. For the phoneme isolation task, which was a free choice task a percentage threshold of 42% was chosen following an inspection of the histogram for phoneme isolation performance. While 86% of the children with low letter knowledge were above chance on the blending task, only 50% and 6 % were above chance on the explicit rhyme and phoneme isolation tasks respectively. When we look at the percentages for rime and phoneme level items separately we can see that the percentage of children with low letter knowledge who succeed on the blending task is lower for phoneme than rime level items but is still high at 72%. On the phoneme isolation task none of the low letter knowledge children succeeded on the phoneme level items.

Table 2.7. Percentage of children within the low letter knowledge group who succeeded on each measure of implicit PR and explicit PR (% success threshold given in brackets)

Task		% children in low letter knowledge group (<3 letters known, N=36) who succeeded		
		All items	Rime level items	Phoneme level items
Implicit PR tasks	Mis. reconstruction (25)	61	56	72
	Pseudoword similarity (50)	61	44	56
	Mis. conflict (25)	69	64	44
	Incomplete word (25)	53	56	36
	Implicit rhyme (25)	44	44	-
Explicit PR tasks	Explicit rhyme (25)	50	50	-
	Blending (25)	86	81	72
	Isolation (42)	6	6	0

Discussion

The aim of the current study was to test predictions made by the key theoretical accounts of phonological development (Table 2.1). To this end a battery of implicit PR measures was developed which allowed us to tap into the segmentedness of children's phonological representations in a way that is separable from the measurement of children's explicit knowledge of the sound system (a construct referred to as phonological awareness). These measures were used alongside traditional explicit PR measures and measures of vocabulary and letter-sound knowledge, allowing us to assess which of the key theoretical accounts is best supported by children's performance.

Implicit PR, explicit PR, vocabulary and letter-sound knowledge were all found to be significantly correlated when controlling for age, confirming that there are important relations between these measures independent of general developmental factors. Evidence was found of emerging PR segmentedness within the youngest nursery group who performed just above chance on implicit PR measures. By the end of reception implicit PR performance had increased, with children choosing the segmental response most of the time. Within the oldest reception group segmental sensitivity still remained significantly below adult levels suggesting that reception children are more heavily influenced by global similarity than adults. This finding is in agreement with Treiman and Breaux's work (Treiman & Breaux, 1982), which shows adults to be more attuned to phonemic versus global similarity than children are (Treiman & Breaux, 1982).

While the accessibility viewpoint proposes that phonological representations are fine grained from infancy (Rozin & Gleitman, 1977; Liberman, Shankweiler & Liberman, 1989), emergent theories suggest that phonological representations only become segmental during childhood following a period of gradual restructuring. Within the emergent account, three key variants have been proposed which differ in terms of the relative roles of vocabulary growth and orthographic knowledge on PR segmentedness.

The role of vocabulary growth

While the accessibility position predicts no unique relation between vocabulary growth and PR segmentedness, emergent theories predict that children's phonological representations will become more componential as they learn more words. The results of the current study support the idea of lexical restructuring with vocabulary being found to predict children's implicit sensitivity to segmental structure while controlling for age. It could be argued that if children's representations were segmented from infancy (as proposed by the accessibility account) we might still expect to find a significant relation between vocabulary and implicit PR measures because of general ability and age related

factors (e.g. children with higher vocabulary scores may perform better on implicit PR measures because they are better at listening, following instructions etc.). However the fact that vocabulary predicts implicit PR performance over and above age, but does not predict PA performance weakens this argument. Furthermore if the accessibility viewpoint held true and children represented words segmentally from infancy, then we would also expect to see much higher levels of performance on the implicit PR measures (which do not require explicit awareness) than those that we observed within the youngest nursery children (only just above chance).

The role of letter-sound knowledge

According to psycholinguistic grain size theory letter-sound knowledge is needed in order for children's phonological representations to become phonemic (Ziegler & Goswami, 2005). Other authors have suggested however that fine grained representation occurs in the absence of letter-sound knowledge, either because it is present from infancy (e.g. Ballem & Plunkett, 2002; Swingley, 2009) or because it emerges gradually through oral language experience under the pressure of a growing vocabulary (Metsala & Walley, 1998, Ventura et al, 2007). This study provides two sources of support for the latter view that letter-sound knowledge is not in fact needed for phonemic representation. Firstly the regression results showed implicit PR not to be predicted by letter-sound knowledge and secondly there were relatively high numbers of children who scored above chance on implicit PR measures despite knowing less than three letters. This number remained high when we looked at phoneme level items, eliminating the possibility that children performed above chance due to success on items requiring onset-rime segmentation only.

Although the theoretical accounts (summarised in Table 2.1) differ in terms of the proposed role of letter-sound knowledge in the development of children's underlying representations, all but the LRM predict a close link between phonological awareness and letter-sound knowledge. This universal prediction is confirmed by the current results, with letter-sound knowledge found to significantly predict explicit PR performance while controlling for age. Inspection of children's performance on the individual explicit tasks when grouped according to letter knowledge suggests that letters may be particularly important for developing phoneme isolation ability but not so important for blending. While the majority of children who knew less than 3 letters could blend phonemes together to make words, only 6% could isolate the onsets of words and none of them could isolate phonemes further on in the word (i.e. the vowel or coda).

The fact that children with very limited knowledge of letters can blend phonemes together may be explained by that fact that blending is arguably the least explicit of the

phonological awareness tasks. Studies show that blending is one of the earliest PA skills for children to master (Anthony, Lonigan, Driscoll, Phillips & Burgess, 2003) and the extent to which children need to explicitly access the phonemes in words to complete this task is unclear. Phoneme isolation on the other hand is arguably the most direct way of assessing children's knowledge of phonemes given that it asks children to tell you explicitly what the phonemes in a given word are. Phoneme isolation may therefore be a 'purer' measure of explicit phonological awareness than blending, which might explain why there is inconsistency across the explicit PR tasks in terms of the performance of the low letter knowledge group.

The overall pattern of results (see Table 2.1) is consistent with Ventura's proposition (Ventura et al, 2007) that while phonological representations may become segmented without the need for letter knowledge, explicit access to the sounds in words for use in phonological awareness tasks requires a grasp of phoneme-grapheme correspondences. Ventura et al.'s study presented evidence of PR segmentation in adults who knew few or no letters (Ventura et al., 2007). This study adds to the literature by replicating this important result within a child sample. Although previous studies have shown evidence of restructuring in children, none (to the authors' knowledge) contrasted implicit and explicit measures of phonological representation alongside letter-sound knowledge and vocabulary. This simultaneous measurement allows a sharper theoretical insight into the processes underlying phonological development at the onset of literacy instruction.

The current study also extends existing work by breaking down the measurement of PR segmentedness into onset-rime versus phoneme level segmentedness – two levels which are proposed to emerge sequentially within the lexical restructuring model (Metsala & Walley, 1998). In this way we have been able to test the prediction made by psycholinguistic grain size theory that phonemes will only emerge within children's representations once letter-sound correspondences have been learnt. We have presented evidence that preliterate children's representations are segmented beyond the onset-rime level – with some children who know less than three letters showing success on measures of phoneme level segmentedness, which cannot be carried out on the basis of comparing onsets or rimes only. The fact that preliterate children can make phonemic classifications when global similarity has been controlled for therefore contradicts Ziegler and Goswami's (2005) proposition that phonological representations only become stored in terms of shared phonemes once children learn the correspondences between letters and sounds.

Conclusion

In conclusion this study has tested the key theoretical accounts of phonological development using novel implicit measures of PR segmentedness alongside traditional measures of explicit phonological awareness. We have presented evidence that while letter-sound knowledge may be important for success on most explicit PR tasks, it may not be required for the representations themselves to become segmented. Evidence was found of children with very limited knowledge of letters succeeding on tasks designed to probe the segmentedness of phonological representations at the phoneme level. The results support the view that lexical restructuring is predominantly driven by oral language experience and that phonemes may emerge within the lexicon in the absence of literacy (Ventura et al., 2007). This study highlights the ability of implicit measures of phonological representation to detect the effects of lexical restructuring independent of orthography.

CHAPTER 3

PHONEMES EMERGE FROM WORDS BUT ARE NOTICED THROUGH LETTERS

Stephanie AINSWORTH

Anne HESKETH

Anna WOOLLAMS

Stephen WELBOURNE

Paper prepared for publication (not yet submitted)

Abstract

Phonological representations and phonological awareness have both been identified as key predictors of literacy success. Yet substantial debate remains about how these entities develop and the relative roles that vocabulary growth and orthographic knowledge play in the development of phonological representation at the onset-rime and phoneme level. The current longitudinal study plots children's performance (across 4 time points) on implicit and explicit measures of phonological representation plus vocabulary and letter-sound knowledge as they move through the first two years of school (N=24, overall age range 3;2 to 5;2). The study provides separate analyses of performance at the rime and phoneme level allowing a more comprehensive test of current theoretical accounts than has been previously possible. The results show that vocabulary is a key predictor of both rime and phoneme level phonological representation on implicit tasks, whereas letter-sound knowledge is a key predictor of phonological representation on explicit tasks at the phoneme level. The results are consistent with the view that while children's phonological representations may be restructured independent of literacy, letter-sound knowledge is needed for children to gain a conscious awareness of phonemes.

Introduction

Phonological representations are the way that children store their knowledge of the sounds in words. There is substantial debate in the literature as to whether young children's phonological representations are adult-like or whether they are qualitatively different, becoming mature only after a protracted period of restructuring. There is also disagreement surrounding the emergence of phonological awareness (the ability to reflect on/and or manipulate phonological segments within words) and the extent to which this is dependent on orthographic knowledge (Castles & Coltheart, 2004). Given that both the quality of phonological representations (measured by mispronunciation detection) and children's explicit phonological awareness of them (measured by rhyme oddity, blending, phoneme deletion etc.) have been shown to predict later reading success (Carroll & Snowling, 2004, Lonigan, Burgess & Anthony, 2000) it is important for us to more fully understand the development of these key competencies.

Theoretical accounts of phonological development

Proponents of the 'accessibility position', propose that children's phonological representations are fine-grained from infancy (e.g. Bailey & Plunkett, 2002, Ballem & Plunkett, 2005, Swingley & Aslin, 2000; Swingley, 2009) but that children are not able to access them explicitly until they develop the required metacognitive skills and orthographic knowledge. It is suggested that children only become conscious of the segmental structure of words once they have grasped the alphabetic principle, i.e. that letters or groups of letters are used to represent segments of sounds (Rozin & Gleitman, 1977; Liberman & Shankweiler & Liberman, 1989). Evidence in support of the accessibility viewpoint comes from studies showing children to be sensitive to phonemic contrasts from as early as 14 months (Bailey & Plunkett, 2002, Ballem & Plunkett, 2005, Swingley & Aslin, 2000; Swingley, 2009; see Ramon-Casas & Bosch, 2014 for a review) and from those which show children's conscious awareness of phonemes to rapidly rise with the onset of literacy instruction (see Ziegler & Goswami, 2005 for a review).

The opposing view, referred to as the emergent position, suggests that children's phonological representations only become fine grained later in childhood following a gradual period of lexical restructuring (e.g. Fowler, 1991; Metsala & Walley, 1998). It is argued that, although infant studies show children to be sensitive to small differences in words, they do not tell us whether or not infant's representations are stored componentially. In other words an infant may be able to tell the difference between two words differing by just one phoneme, but that does not necessarily mean that they store phonological forms in terms of phonemes, e.g. storing the /b/ in 'bed' in the same way as

the /b/ in 'cab' (Ziegler & Goswami, 2005). Within the emergent view early phonological representations initially lack detail and become gradually augmented as children's lexicons grow. There are three key variants of the emergent view, outlined below, which all share the idea of vocabulary driven restructuring taking place over childhood but differ in terms of the proposed influence of literacy instruction on the way that children store phonological forms.

The lexical restructuring model or LRM (Metsala & Walley, 1998) suggests that infants' phonological representations are stored holistically, i.e. as a whole word template or in terms of one particularly salient feature within the word. The idea here is that when children's lexicons are small, they do not need to store words in much detail in order to keep them distinct. As they learn more words however, and the lexicon becomes populated with phonologically similar words (e.g. cat-hat and cat-cot), children need to restructure their representations in order to avoid confusion and keep representations distinct. The LRM proposes that this process involves a developmental increase in the segmentedness of children's phonological representations with words stored initially as whole word forms, then as onset and rime segments (c-at) and finally as phonemic segments (c-a-t). This process is proposed to take place on a gradual word by word basis with the most frequent words residing in the densest neighbourhoods being restructured first (Metsala & Walley, 1998; Walley, Metsala & Garlock, 2003). Within the LRM it is suggested that the level of segmentedness of children's phonological representations sets the stage for children's later conscious awareness of the sound segments within words. In other words, before children can explicitly access rime segments (e.g. on rhyme awareness tasks) their representations first need to be segmented at least at the onset-rime level. Similarly, to access phonemes explicitly on phoneme awareness tasks children first need to have representations which are segmented at the phoneme level.

Support for the LRM comes from studies which show children's early representations to lack detail (Altvater-Mackensen, van der Feest & Fikkert, 2013) and from studies which show children to perceive words holistically in contrast to adults who attend to the segments within words (see Metsala & Walley, 1998 for a review). For example in mispronunciation detection tasks, while adults were found to detect mispronunciations more easily in initial position, children's ability to detect mispronunciations was found to be less dependent on the position within the word (Cole & Perfetti, 1980; Walley, 1987). Bowey & Hirakis (2006) however have questioned the use of position effects as a reliable test of the restructuring hypothesis, based on their finding that position effects covary with the clarity of acoustic phonetic information in both children and adults. Bowey and Hirakis (2006) found that when the potential confound of

clarity was controlled for (by presenting words where the syllable is stressed or the vowel is not reduced) both children and adults detected mispronunciations in the first and second syllable equally well. Further support for the lexical restructuring model comes from an interaction between frequency and neighbourhood density on a range of speech perception tasks (Metsala, 1997; Garlock, Walley & Metsala, 2001) consistent with the idea of high frequency words from dense words being restructured first (Metsala & Walley, 1998).

While the lexical restructuring model acknowledges that literacy may play a role in the development of children's phonological representations, the model focuses on vocabulary growth as the key driver of lexical restructuring (Metsala & Walley, 1998). Psycholinguistic grain size theory (Ziegler & Goswami, 2005) on the other hand affords literacy instruction a more central role, suggesting that phonological representations will only become structured in terms of phonemes once children learn about phonemes explicitly in school. It is suggested that the most common way for children to become aware of phonemes is through being taught the correspondences between letters and sounds (or more generally the mappings between graphemes – which can include letter strings of more than one letter, e.g. *ch* – and phonemes). Ziegler and Goswami (2005) suggest that while rime segments will emerge naturally within the lexicon through oral language experience alone (due to the salience of rhyme within the English language) phonemic structure will only appear once children are taught about phonemes and/or the mappings between letters and sounds. Support for the importance of orthographic knowledge on the emergence of phonemic representation comes from studies which show a steep rise in phoneme awareness ability at the time when children start learning to read (see Ziegler & Goswami, 2005 for a review).

The final variant of the emergent view was put forward by Ventura and colleagues based on their work with adult groups with varying levels of literacy (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007). Ventura et al. (2007) found evidence of lexical restructuring in the absence of letter knowledge with adult illiterates showing the interaction between neighbourhood density and frequency predicted by the lexical restructuring model (Metsala & Walley, 1998). Ventura and colleagues extended the scope of the lexical restructuring model by proposing that while letter-sound knowledge is not required for the restructuring of the representations themselves, it is important for gaining conscious access to phonological segments during explicit phoneme awareness tasks (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007). This argument draws on numerous studies which have shown that while illiterate adults perform in line with their literate peers on explicit rhyme awareness tasks (where they are asked questions such as, 'Do *dak* and *lak* rhyme?', de Gelder, Vroomen & Bertelson, 1993), illiterates perform

poorly on tasks which measure explicit phonological awareness at the level of individual phonemes (e.g. phoneme deletion tasks – ‘When I say *kur* you say *ur*’, de Gelder, Vroomen & Bertelson, 1993).

Predictions made by key theories of phonological development

The accessibility view and the three variants of the emergent view each make different predictions about the relative roles of vocabulary growth and letter-sound knowledge on the development of phonological representation (PR) and explicit phonological awareness (PA) as summarised in Table 3.1. Each account predicts different causal pathways for phonological development at the rime and phoneme levels (with the exception of the accessibility view which does not make any specific predictions about rime level development) as illustrated in Figure 3.1. Each box in the figure denotes an aspect of the child’s knowledge, which we aimed to assess with a set of age appropriate behavioural measures (described within the method section of this chapter).

Within the accessibility account, neither vocabulary growth or letter-sound knowledge are predicted to be key drivers of phonological representation given that children’s phonological representations are proposed to be adult-like from infancy. Letter-sound knowledge is however predicted to drive explicit awareness of phonemes given that mastery of the alphabetic principle is suggested to be the key to gaining explicit access to phonemes (Rozin & Gleitman, 1977; Liberman & Shankweiler & Liberman, 1989).

Turning next to the three variants of the emergent account we can see that they all agree that both rime level representation and rhyme awareness emerge naturally through oral language experience alone. Where the theories differ is in their predictions about the emergence of the phoneme (see the bottom set of pathways within Figure 3.1). Within psycholinguistic grain size theory it is suggested that words are not represented in terms of phonemes until children develop an explicit awareness of phonemes when learning letter-sound mappings (Ziegler & Goswami, 2005). Psycholinguistic grain size theory therefore predicts that letter-sound knowledge will be the key driver of representation at the phoneme level. Ventura and colleagues’ account on the other hand suggests that restructuring is independent of literacy with vocabulary driving the emergence of segmental representation (Ventura et al, 2007). Although the lexical restructuring model (Metsala & Walley, 1998) makes no specific predictions about the role of letter-sound knowledge, given the model’s emphasis on lexical growth as the force behind PR segmentation it is assumed that vocabulary growth would be predicted to have a greater influence than letter-sound knowledge within this model.

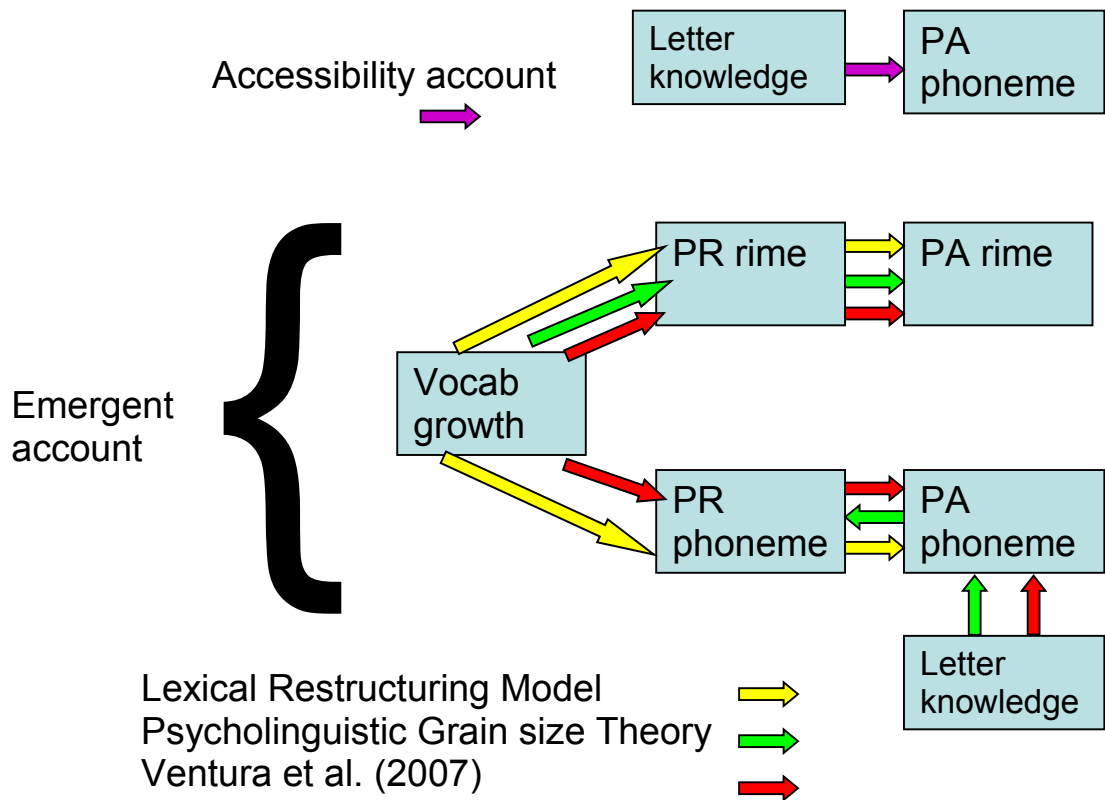


Figure 3.1. Causal pathways predicted by four theoretical accounts of phonological development

The final set of predictions made by the different variants of the emergent view relate to the emergence of children’s explicit phoneme awareness. From the final column of Table 3.1 we can see that both psycholinguistic grain size theory (Ziegler & Goswami, 2005) and Ventura et al.’s account (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007) both share the accessibility view’s prediction that children’s conscious access to phonemes will be strongly related to their letter-sound knowledge. This is because each theory sees mastery of the alphabetic principle as the most common way that children’s attention is drawn explicitly to phonemic segments. No specific prediction is given within the lexical restructuring model relating to the role of letter-sound knowledge and so an ‘unspecified’ entry has been given. Note that within Figure 3.1 there is not only an arrow feeding into phoneme awareness letter knowledge but also an arrow feeding in from PR phoneme. This is because Ventura and colleagues found evidence that the development of phoneme awareness is constrained by lexical restructuring (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007). Adult performance on a phonological awareness task (phoneme deletion) was found to be affected by the neighbourhood density of the word items. Ventura and colleagues concluded that while letter knowledge precipitates the development of phonological awareness, the degree to which the representations have been

restructured constrains PA performance (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007).

The predictions made by the theories of phonological development listed above involve four key constructs: the level of segmentedness of children’s representations, children’s explicit awareness of the sounds stored within their representations, letter-sound knowledge and vocabulary size. In order to test these predictions we might attempt to measure each of these constructs in turn and examine the relations between them. While the measurement of the latter three constructs is relatively straightforward (e.g. when measuring explicit phonological awareness we can ask children ‘What are the sounds in ‘cat’? Which one rhymes with ‘sun’?’ etc.), we are not able to measure children’s phonological representations directly. We can however attempt to probe the segmentedness of children’s representations using implicit tasks which require children to make judgements about phonological word forms but do not require children to have any explicit knowledge of the sound structure of words (Ainsworth, Welbourne, Woollams & Hesketh, submitted; Storkel, 2002; Carroll & Myers, 2011).

Table 3.1. A summary of the relative roles of vocabulary growth and letter-sound knowledge on phonological representation and phonological awareness at the rime and phoneme level.

	Key driver of...			
	Rime level		Phoneme level	
	Representation (PR rime)	Awareness (PA rime)	Representation (PR phoneme)	Awareness (PA phoneme)
Accessibility view ^a	Neither	Unspecified	Neither	Letter-sound knowledge
Lexical restructuring model ^b	Vocabulary growth	Vocabulary growth	Vocabulary growth	Unspecified
Psycholinguistic grain size theory ^c	Vocabulary growth	Vocabulary growth	Letter-sound knowledge	Letter-sound knowledge
Restructuring but not phoneme awareness occurs in the absence of literacy ^d	Vocabulary growth	Vocabulary growth	Vocabulary growth	Letter-sound knowledge

^a e.g. Rozin & Gleitman (1977), Liberman, Shankweiler & Liberman (1989), ^b Metsala & Walley (1998), ^c Ziegler & Goswami (2005), ^d Ventura, Kolinsky, Fernandes, Querido & Morais (2007)

A recent cross-sectional study used both implicit and explicit measures of phonological representation within a sample of nursery and reception children to test the predictions set out in Table 2.1 (Chapter 2). The results found a dissociation between performance on implicit versus explicit measures of phonological representation with vocabulary growth predicting implicit PR performance and letter-sound knowledge predicting explicit PR performance. These results are consistent with Ventura and colleagues' proposition that phonological representations may be restructured without literacy but that letter knowledge is needed for explicit access to phonemic segments on phoneme awareness tasks (Ventura et al, 2007). The results were based on children's overall performance where scores on items measuring rime versus phoneme level representation were not separated (although Chapter 2 looked descriptively at performance on rime versus phoneme level items, the study had insufficient power to perform the main analysis on the two levels of segmentation separately). We have seen above (Figure 3.1 and Table 3.1) however, that the four theoretical accounts make important distinctions between the development of phonological representation and phonological awareness at these two levels of segmentation. In order to make a more comprehensive evaluation of these accounts we need to investigate the relative roles of vocabulary growth and letter-sound knowledge on the development of rime and phoneme level representation and awareness separately.

The current study aims to plot the development trajectory of phonological knowledge as children move through the first two years of school (aged 3;2 to 5;7). A combination of implicit and explicit measures of phonological representation are used allowing us to probe both the segmentedness of the representations themselves as well as children's conscious awareness of phonological segments. The longitudinal design extends previous work by providing the required power to analyse the development of onset-rime versus phoneme level representation separately. This in turn allows us to test predictions from theory at a greater level of detail than has previously been possible.

Method

Participants

24 children were recruited from the nursery classes of two UK mainstream primary schools of low and medium socio-demographic status respectively, as indexed by the percentage of free school meals (based on the 2012 school dashboard data, see Ofsted, 2014). Ethical approval for the study was granted from the University Ethics Committee. The sample consisted of the youngest 24 children from a group of 90 recruited for an earlier cross-sectional study using the same measures (Chapter 2). Within the current study the young nursery group (aged 3;2 to 3;10) were followed up on three occasions roughly five months apart, in addition to the initial testing point which was conducted in the Autumn term of the nursery year as part of the earlier cross-sectional study. The gender balance was 15 boys and 9 girls. To be included in the study children needed to have at least one English speaking parent, to have no known history of speech or hearing problems (as reported by the teacher) and to not be on the special educational needs register for any behavioural or developmental concerns. All children spoke fluent English. We do not have complete information about the children's home language environments, although for the majority of children within the study English was the child's first language.

Procedure

Children were tested in a quiet room within school, over 5 to 10 sessions depending on their age/attention span. All expressive tasks were audio-recorded for checking purposes. Corrective feedback was given for one training item at the beginning of each task (with the exception of the expressive vocabulary and letter sound knowledge tests which did not have any training items). For all other items general praise and stickers were given as encouragement independent of performance.

Materials

Items were selected as being familiar to young children – 98% of the words used can be found within Storkel and Hoover's database (Storkel & Hoover, 2010) which is based on corpora of kindergarten and first grade children. At the beginning of each trial children were asked to name the picture stimuli and were told the word if they were unable to identify the picture correctly.

All the distracters within the multiple choice implicit and explicit PR tasks were matched listwise in terms of frequency and two measures of phonotactic probability (positional segment average – which measures how often the segments within a word occur in that position within other words and biphone average which measures the

frequency of pairs of sound segments) using Storkel and Hoover's online calculator (Storkel & Hoover, 2010). See Appendices 1 to 7 for stimuli lists and Appendices 8 to 10 for additional details on matching.

Implicit measures of phonological representation

Children were tested on five implicit measures of phonological representation. While four of the tasks probe the segmentedness of children's representations at both the onset-rime and phoneme level, one task probes segmentedness at the onset-rime level only. A summary table of all tasks and illustrated examples are given in Chapter 2. For the sake of brevity the tasks measuring children's implicit knowledge of their phonological representations will henceforth be referred to as implicit PR tasks.

Similarity based classification tasks

Three of the implicit tasks ask children to make similarity based classifications: mispronunciation reconstruction, pseudoword similarity and mispronunciation conflict. In each case children hear a CVC stimulus and are asked to choose which word/pseudoword it sounds the most like out of two or more choices. For each trial the typical adult response (see Chapter 2) is a word/pseudoword which shares two phonemes with the stimulus (e.g. *nake*-name) and the closest distracter is a word/pseudoword which only shares one phoneme with the stimulus but is matched in terms of global similarity (e.g. *nake*-net). Further details on global similarity matching are given below. The rationale behind the three similarity based implicit PR tasks is that if children choose the closest segmental match more often than the globally matched distracter, then we can infer that they are sensitive to the number of shared segments over and above how similar the words/pseudowords are in terms of global similarity (overall 'sounds-likeness').

Sensitivity to shared segments in turn may be interpreted as evidence that the representations are segmented at least at the grain size of those segments, i.e. if children are sensitive to shared rimes then we can infer that their phonological representations must be segmented at least at the onset-rime level. Similarly if children are sensitive to shared phonemes this suggests that they are storing words phonemically. All three tasks contain items designed to probe the segmentedness of children's representations at both the rime and phoneme levels. While rime level items contain two shared phonemes in rime position (e.g. *ven*-ten), phoneme level items have two shared phonemes in body position (e.g. *shet*-shell). It is assumed that while rime items can be completed if the child's phonological representations are segmented at the rime level only (by comparing the rime segments as a whole), a more fine grained phonemic representation is needed for phoneme level items

where the rime segment is disrupted (see Metsala & Walley, 1998 and Ziegler & Goswami, 2005 for a review of evidence in support of primacy of the rime unit in English).

Global similarity matching

For each of the similarity-based measures the closest distracter (or only distracter in the case of the two-choice tasks) was matched to the segmental response in terms of global similarity. For example when children are asked to choose whether ‘rain’ or ‘pin’ sounds the most like ‘*hain*’ the distracter ‘pin’ is just as close to ‘*hain*’ in terms of global similarity as ‘rain’ despite sharing only one (rather than two) phonemes with ‘*hain*’. Global similarity scores were calculated using adult ratings collected by Singh and colleagues (Singh & Woods, 1971; Singh, Woods & Becker, 1972). Scores of how dissimilar word/pseudoword stimuli were from one another were calculated using the same additive method adopted by Treiman and Breaux (1982) and Carroll and Snowling (2001). For example the dissimilarity score between the words ‘pin’ (pronounced /pɪn/) and ‘bed’ (pronounced /bed/) is the dissimilarity of /p/ and /b/ (3.9) plus the dissimilarity of /ɪ/ and /ɛ/ (2.22) plus the dissimilarity of /n/ and /d/ (4.8).

The details of the individual similarity based tasks are as follows:

Mispronunciation reconstruction Children heard a puppet mispronounce a word and were then asked to guess which picture he was trying to say – which picture did it sound the most like. For example the puppet said ‘*hain*’ and the children chose whether he was trying to say rain, pin, bone or tap.

Pseudoword similarity Children were asked which of two pseudoword alien names sounded the most like a third pseudoword name. For example, ‘Which one sounds the most like *Ses*, is it *Sen* or *Sif*?’

Mispronunciation conflict Children heard three aliens attempt to say a word (with the target picture shown on screen) and were first asked to choose which one said the word correctly. This part of the task was included to check that the children were able to distinguish the correct form of the word from the mispronunciations. Children then listened to the two mispronunciations again and were asked, ‘Which one said it the best? Which one sounded the most like it?’ For example children chose whether ‘*yail*’ or ‘*pell*’ sounds most like ‘tail’.

There were two further implicit PR tasks:

Incomplete word Children heard a word onset and were asked which picture (out of a choice of four) the puppet wanted, where the target picture was of a word sharing the same onset spoken by the puppet. For rime level items, the onset was a single consonant (e.g. /k/ with target picture ‘cup’). For phoneme level items children were required to link a consonant cluster to a word (e.g. /fl/ - flag’), where the closest distracter began with a cluster sharing the same initial consonant (e.g. ‘frog’). It is assumed that while children can link /k/ to ‘cup’ when their phonological representation for ‘cup’ is segmented at the onset rime level only, children will need to have a finer grain phonemic representation to be able to distinguish between the /fl/ in ‘flag’ and the /fr/ in ‘frog’. As with all the implicit PR tasks the incomplete word task is designed to be as implicit as possible: children are not required to explicitly reflect on the sounds in words (as they are in traditional PA tasks, e.g. ‘What is the first sound in cup?’), rather they are just asked to guess what the puppet wants, having heard it say a sound.

Implicit rhyme Children heard a puppet say three rhyming words (e.g. fed, bed, head) and were asked to choose which picture completes the pattern (e.g. red). In contrast with the explicit rhyme task described below, this task avoids the use of the word ‘rhyme’ and does not ask children to explicitly reflect on the shared rime segments. Rather children are simply required to pick up on the rhyme pattern implicitly.

Explicit measures of phonological representation

Children were tested on three explicit measures of phonological representation (explicit PR). These tasks are designed to measure children’s explicit phonological knowledge. Such measures are often referred to as phonological awareness tasks. We use the term explicit PR within the current paper to emphasise that the test battery as a whole seeks to measure two aspect of phonological representation: children’s implicit sensitivity to the phonological segments within words (implicit PR) and children’s conscious awareness of the phonological segments within words (explicit PR).

Explicit rhyme Children were shown four pictures (e.g. sun, rain, pan, coat) and asked which one rhymes with a given word (e.g. ‘Which one rhymes with fun?’). The same stimuli, target items and distracters used in the implicit rhyme task were also used here to allow a direct comparison across implicit and explicit measures of phonological representation at the same level of segmentedness (i.e. the onset-rime level). The explicit

rhyme task always took place after the implicit rhyme task with a gap of a few days in between.

Blending Children heard a robot (whose voice was pre-recorded) say either an onset and a rime (e.g. c-at) or three individual phonemes (e.g. h-a-t) and were asked to select the corresponding picture.

Phoneme isolation Children were shown a picture and asked to say the sounds in the word (e.g. /m-ʌʊ-s/ if shown a picture of a mouse). To avoid unnecessary testing, the task ended if children were unable to isolate any of the sounds in eight consecutive words, and it was assumed that children would have scored zero on all remaining items (this criterion applied to 18 out of 92 cases).

Background language measures

Letter-sound knowledge Children were shown a grapheme (a letter or group of letters) and were asked ‘What sound does it make?’ The graphemes were presented in order of difficulty as indexed by the order in which the Letters and Sounds framework (PNS, 2007) recommends that they are taught. Testing was stopped after 8 consecutive incorrect responses and the child was scored as knowing the number of letters answered correctly up to this point (this criterion applied to 57 out of 92 cases). It should be noted that within this paper we use the term ‘letter-sound knowledge’ broadly to refer to all grapheme-phoneme correspondences (e.g. including the mappings for digraphs like ‘sh’ – /ʃ/), not just to the mappings between single letters and sounds.

Vocabulary Expressive and receptive vocabulary were measured using the Renfrew Word Finding Vocabulary Test (Renfrew, 1995) and British Picture Vocabulary Scale III, BPVS (Dunn, Dunn, Sewell, Styles, Bryzaska, Shamsan & Burge, 2009) respectively. Standardised procedure for the BPVS was adapted slightly: all children started at set 1 regardless of age so that raw scores could be compared between age groups.

Counterbalancing

For all of the implicit and explicit PR measures, two sets of item orders were used to control for order effects. Children were randomly assigned to either set at time point 1. The same set was then used at all subsequent time points. The order in which the tasks were delivered was held constant across participants and time points. This was done to ensure that the explicit tasks always followed the implicit tasks. If children carried out the

explicit tasks first, this may have impacted on their implicit performance due to increased salience of the segments accessed within the explicit tasks. Given that we were interested in separating children's implicit and explicit knowledge as much as possible, it therefore made sense to conduct the explicit measures after all the implicit measures had been completed.

Results

Data screening

Twenty out of the twenty four children completed all tasks at all time points. A summary of the missing data from the remaining four children is given in Table 3.2. All data points were used in the analyses that follow unless otherwise stated. The box plots for all tasks were inspected for the presence of outliers – none were found.

Table 3.2. Number of tasks completed at each time point by the four children with incomplete data profiles (max 11)

	Time point 1	Time point 2	Time point 3	Time point 4
Child 1	11	11	0	0
Child 2	4	11	11	11
Child 3	11	11	11	8
Child 4	11	11	11	10

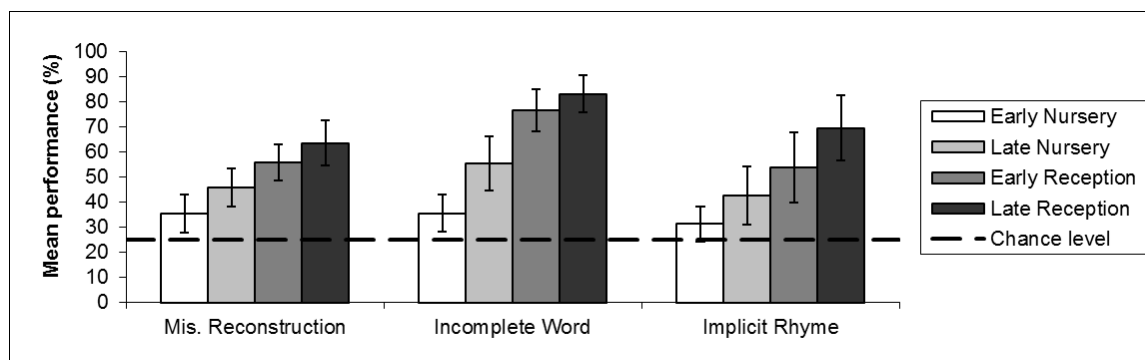
Developmental performance on implicit and explicit measures of phonological representation

The mean percentage performance on each of the five implicit PR and 3 explicit PR tasks are shown in Figures 3.2 and 3.3. At the beginning of nursery (aged 3;2 to 3;10) children were significantly above chance on 3 out of 5 of the implicit measures of phonological representation as confirmed by a series of one sample t tests: mispronunciation conflict ($t(23)=3.25, p=.004$), mispronunciation reconstruction ($t(23)=2.74, p=.012$) and incomplete word ($t(23)=2.76, p=.011$). Children did not perform significantly above chance on the pseudoword similarity and implicit rhyme tasks until time point 2 (aged 3;7 to 4;3, pseudoword similarity: $t(23)=4.21, p<.0005$, implicit rhyme: $t(23)=2.99, p=.007$). Repeated measures ANOVA showed a significant effect of time point for 4 out of the 5 tasks: mispronunciation reconstruction ($F(3,66)=16.58, p<.0005$), mispronunciation conflict ($F(3,66)=7.55, p<.0005$), incomplete word ($F(3,66)=49.98, p<.0005$) and implicit rhyme ($F(3,66)=11.35, p<.0005$). Figure 3.2b shows a small increase in performance on the pseudoword task with each time point but this did not quite reach significance, $F(3,63)=2.63, p=.058$).

On the explicit PR measures children were significantly above chance on the blending task ($t(22)=4.12, p<.0005$) at the beginning of nursery but were not above chance

on the explicit rhyme task ($t(23)=1.62, p=.11$). On the phoneme isolation task (a free choice task with no chance level) children performed poorly at time point 1 (mean=13%, SD=13%) and did not get more than half of the responses correct until the beginning of reception (time point 3: mean=75%, SD=18%). Repeated measures ANOVA showed performance on all three explicit PR tasks to increase significantly over time (blending: $F(3,60)=32.57, p<.0005$; explicit rhyme: $F(3,66)=15.25, p<.0005$; phoneme isolation: $F(3,63)=56.15, p<.0005$).

a)



b)

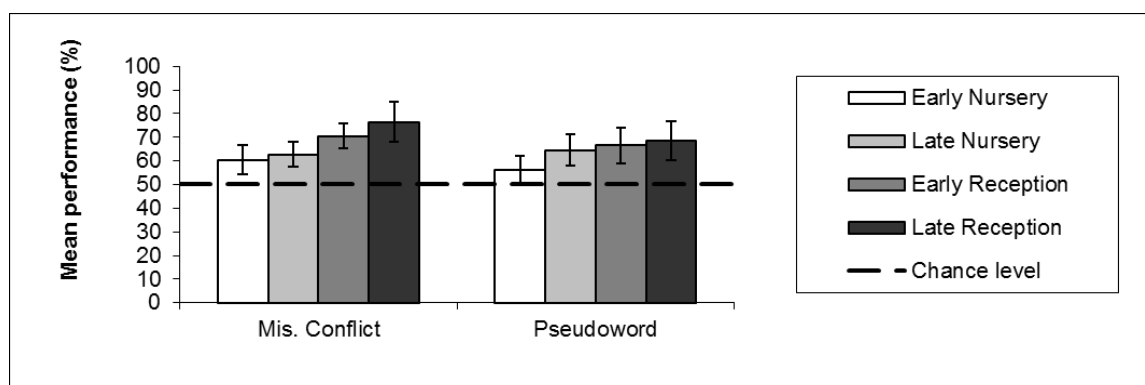


Figure 3.2. Performance on implicit measures of phonological representation – mispronunciation reconstruction, incomplete word, implicit rhyme, mispronunciation conflict and pseudoword similarity – at each time point with a chance level of a) 25% and b) 50%. Mean performance is based on % segmental responses except for the incomplete word scores which are based on % correct. Error bars represent confidence intervals of 95%

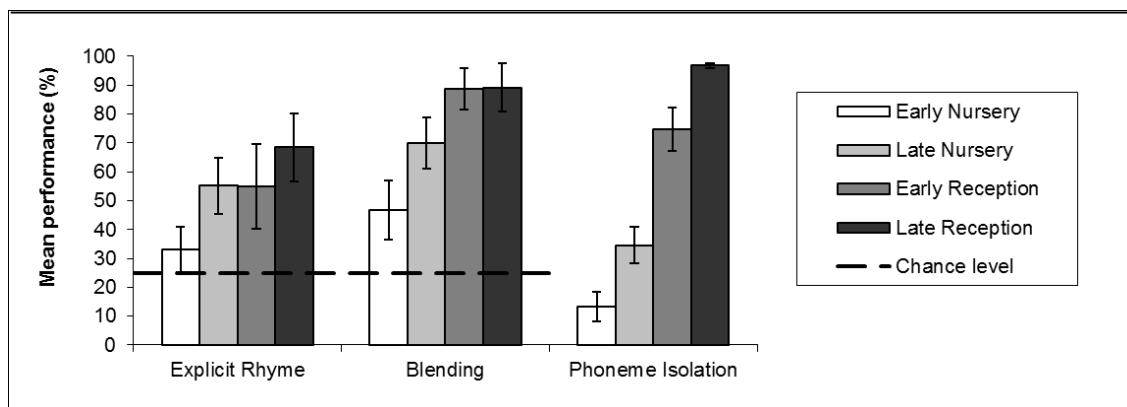
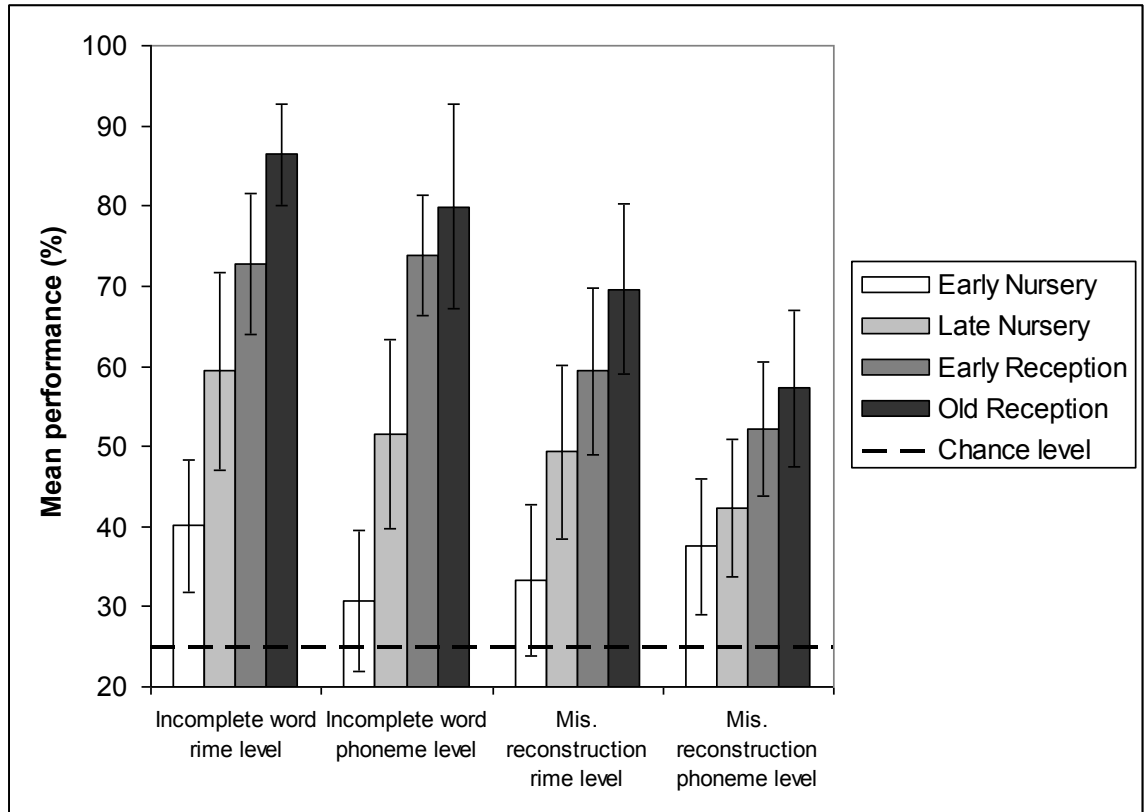


Figure 3.3. Performance on explicit measures of phonological representation. Error bars represent confidence intervals of 95%

Performance on measures of rime versus phoneme level representation

For each task percentage scores were calculated based on the number of segmental responses (or correct responses in tasks where there is a right answer) to rime level items only and the number of segmental (or correct) responses to phoneme level items (Figures 3.4 and 3.5). We can see from Figures 3.4 and 3.5 that phoneme level performance tends to lag behind rime level performance on both implicit and explicit PR measures. A series of paired sample t-tests (with data across all time points entered at once) found children to perform better on rime versus phoneme level items for 3 out of the 4 implicit tasks: mispronunciation conflict ($t(93)=3.73, p<.005$), mispronunciation reconstruction ($t(93)=1.93, p=.03$) and incomplete word ($t(93)=3.60, p=.001$). The effect of grain size was not significant for pseudoword similarity ($t(92)=0.96, p=.34$) but was significant for both explicit measures: blending ($t(91)=3.31, p=.001$) and phoneme isolation ($t(91)=5.77, p<.0005$).

a)



b)

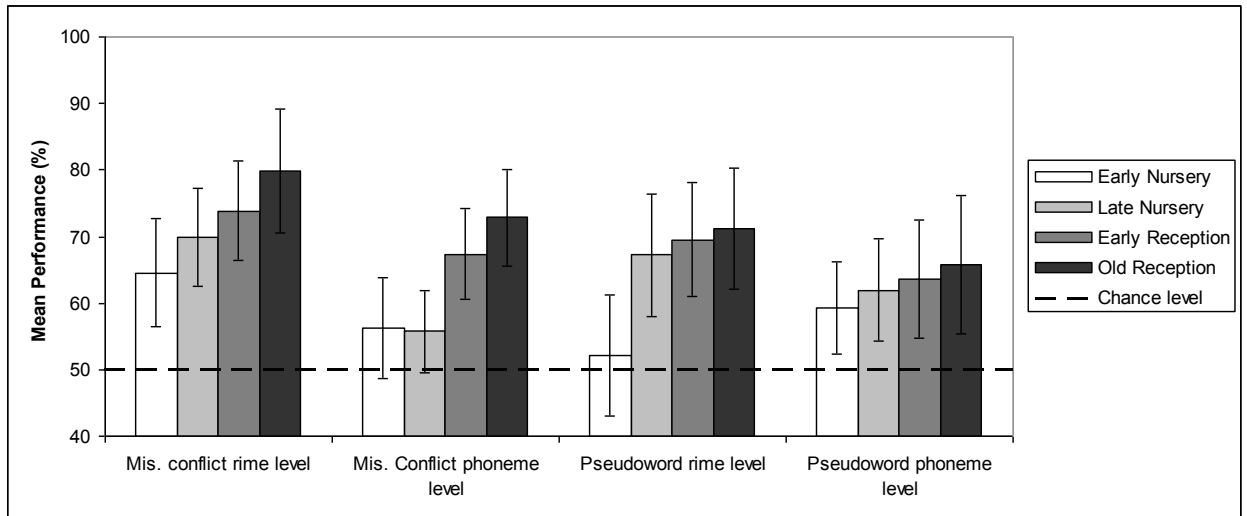


Figure 3.4. Performance on implicit PR measures – incomplete word, mispronunciation reconstruction, mispronunciation conflict and pseudoword similarity with a chance level of a) 25% and b) 50%. Scores have been split into rime level versus phoneme level items. Error bars represent confidence intervals of 95%.

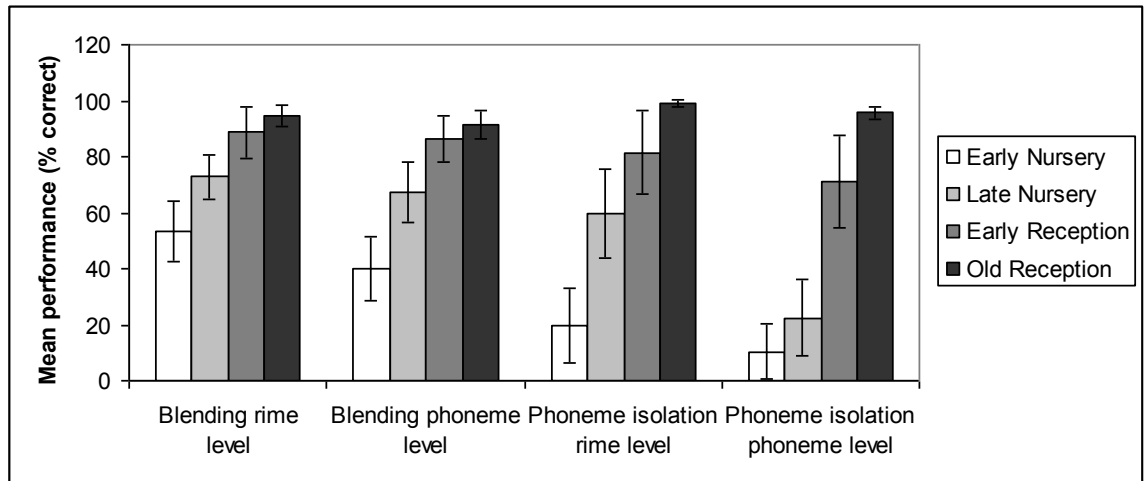


Figure 3.5. Performance on explicit PR measures split into rime level versus phoneme level items. Error bars represent confidence intervals of 95%.

In order to test the predictions set out in Table 3.1 we first reduced the battery of 5 implicit PR measures and 4 explicit PR measures down to four components using principal components analysis. Four PCA analyses were conducted to create weighted averages for implicit PR performance at the rime level, implicit PR at the phoneme level, explicit PR at the rime level and explicit PR at the phoneme level. The first component ‘implicit PR rime’ was extracted from the scores on the rime level items of each of the implicit PR measures and the ‘implicit PR phoneme’ component was extracted from the scores on the phoneme level items for each of the implicit PR measures. Similarly, an ‘explicit PR rime’ and an ‘explicit PR phoneme’ component were extracted from the rime and phoneme level explicit PR scores respectively. We also reduced the two vocabulary measures into one ‘vocabulary size’ component using PCA to create a weighted average of the two measures. Letter-sound knowledge consisted of a single measure and so remained as a stand-alone variable. The component loadings for each task are presented in Table 3.3.

We can see from Figure 3.6 that all six variables (the five principal components plus letter-sound knowledge) increased steadily over time as we would expect (implicit PR rime: $F(3,63)=29.26, p<.0005$; implicit PR phoneme: $F(3,63)=19.84, p<.0005$; explicit PR rime: $F(3,60)=40.26, p<.0005$); explicit PR phoneme: $F(3,60)=52.78, p<.0005$). A high level of interrelatedness was found between the variables even when controlling for age (Table 3.4) with all but one pairwise correlation achieving significance (the relationship between implicit PR phoneme and letter-sound knowledge).

Table 3.3. Task loadings for the six principal components

Task	Implicit PR Rime Component	Implicit PR Phoneme Component
Mispronunciation Reconstruction	.87	.74
Pseudoword Similarity	.72	.67
Mispronunciation Conflict	.73	.73
Incomplete Word	.72	.70
Implicit Rhyme	.83	N/A
R ²	.60	.50
	Explicit PR Rime Component	Explicit PR Phoneme Component
Blending	.84	.91
Explicit Rhyme	.73	.91
Phoneme Isolation	.90	N/A
R ²	.68	.83
	Vocabulary Component	
Expressive Vocabulary (Word Finding Scale)	.96	
Receptive Vocabulary (BPVS)	.96	
R ²	.92	

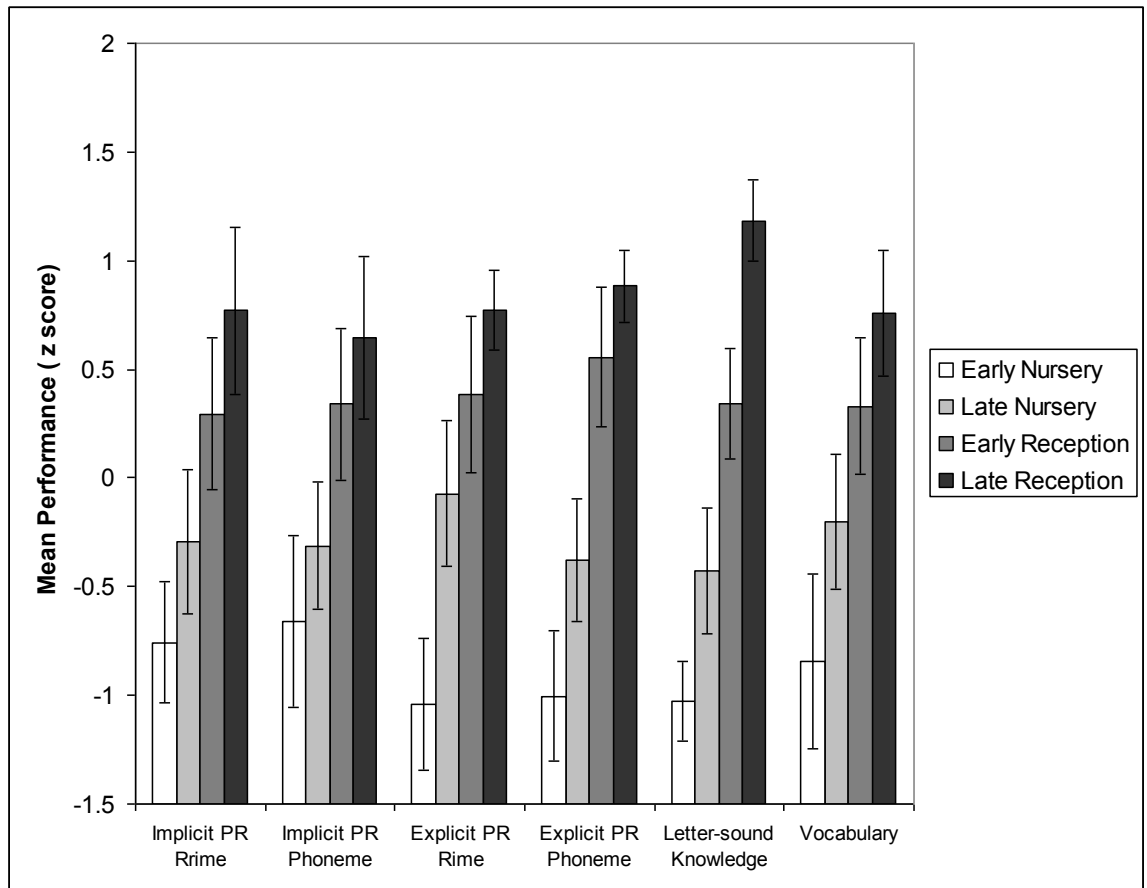


Figure 3.6. Developmental progression across the six variables of theoretical interest. Error bars represent confidence intervals of 95%. The z scores for implicit PR, explicit PR and vocabulary are the factor scores arising from the PCA analysis. The letter-sound knowledge z scores were converted from the raw scores against a normal distribution.

Table 3.4. Partial correlations between the six key variables while controlling for age.

	Implicit PR Rime	Implicit PR Phoneme	Explicit PR Rime	Explicit PR Phoneme	Vocabulary
Implicit PR	.51*				
Phoneme					
Explicit PR	.64*	.53*			
Rime					
Explicit PR	.55*	.40*	.72*		
Phoneme					
Vocabulary	.50*	.41*	.56*	.43*	
Letter- sound knowledge	.45*	.14 (.09)	.38*	.51*	.52*

*p < .0005 (1-tailed)

The relative role of vocabulary growth and letter-sound knowledge in the development of PR segmentedness and explicit phonological awareness

The four key theoretical accounts make differing predictions about the relative importance of vocabulary growth and letter-sound knowledge for the development of children's phonological representations and their explicit awareness of the sound segments within them (Table 3.1). In order to test these predictions, generalised estimating equations (GEEs) were used to test whether vocabulary growth or letter-sound knowledge is more predictive of implicit and explicit measures of phonological representation. Within Figure 3.1 the predictions are visualised in terms of separate pathways for rime level and phoneme level representation respectively. For this reason GEE analysis was performed separately for rime versus phoneme level components leading to four different analyses with implicit PR rime, implicit PR phoneme, explicit PR rime and explicit PR phoneme respectively as the outcome variable (see Table 3.5). In each case the other key variables were included as predictors along with age. Where applicable only predictors of the same grain size (i.e. rime or phoneme) were included, (e.g. when implicit PR rime was the outcome we included explicit PR rime, vocabulary, letter-sound knowledge and age as predictors).

This is because while some of the theoretical accounts predict a relationship between PR segmentedness and phonological awareness (for which our implicit and explicit PR measures are designed to probe respectively) at the same grain size – e.g. the lexical restructuring model predicts that rime level segmentation will set the stage for rime level awareness – none of the accounts make predictions about relationships across grain sizes. For each model the autoregressive working correlation matrix was used given that in a developmental design we would expect measurements at a given time point to predict measurements at subsequent time points. The autoregressive working correlation matrix assumes decreasing correlation for farther time points, e.g. if the correlation between scores at T1 and T2 is ρ then the correlation between scores at T2-T3 and T3-T4 would be ρ^2 and ρ^3 respectively.

The results (listed within Table 3.5 and summarised in Figure 3.7) show that while vocabulary growth is a stronger predictor than letter-sound knowledge of implicit PR rime, implicit PR phoneme and explicit PR rime, letter-sound knowledge is a better predictor than vocabulary growth of explicit PR at the phoneme level. Given that implicit PR and explicit PR measures are assumed to be a proxy for PR segmentedness and phonological awareness respectively, the results indicate that vocabulary growth seems to have a stronger influence than letter-sound knowledge on PR segmentedness at both grain sizes and on rhyme awareness, whereas letter-sound knowledge seems to be important for the emergence of phoneme awareness.

The GEE method was used in the above analyses given its ability to account for the relationships between measures taken at different time points within a longitudinal design (Ballinger, 2004). It is noted that an alternative approach would have been to use a growth curve method. This would have potentially provided a more sophisticated tool for examining the development of individual trajectories over time (e.g. see Duncan & Duncan, 2004). However, within the current study, the limited sample size did not provide sufficient power to allow us to use a growth curve methodology. The authors acknowledge that (as discussed within Chapter 2) the strong relationship found between implicit and explicit PR measures found within both the cross sectional (Chapter 2) and longitudinal data (reported within the current chapter) could have potentially have diluted the apparent influence of letter-sound knowledge on implicit PR. We cannot therefore state with certainty that letter-sound knowledge plays no role in the development of segmental phonology at an implicit level. However, what is clear is that letter-sound knowledge has a greater impact on explicit PR performance at the phoneme level than on implicit PR performance (at both rime and phoneme levels) and explicit PR performance at the rime level.

Table 3.5. GEE results for model 1 with implicit PR rime as the outcome variable and explicit PR rime, vocabulary, letter-sound knowledge and age as predictors

Parameter	B	S.E.	Wald Chi-Square	Sig.
(Intercept)	-1.71	1.12	2.31	.13
Explicit PR Rime	.35	.03	14.23	<.0005
Vocabulary	.23	.10	4.91	.027
Letter-sound knowledge	.011	.014	0.58	.43
Age	.001	.0008	1.40	.23
(Scale)	.42			
Goodness of fit statistics^a	QIC=51.29		QICC=46.39	

^a Computed using the full log quasi-likelihood function, in ‘small-is-better’ form

Table 3.6. GEE results for model 2 with implicit PR phoneme as the outcome variable and explicit PR phoneme, vocabulary, letter-sound knowledge and age as predictors

Parameter	B	S.E.	Wald Chi-Square	Sig.
(Intercept)	-3.33	1.17	8.03	.005
Explicit PR Phoneme	.43	.11	14.74	<.0005
Vocabulary	.33	.095	11.26	.001
Letter-sound knowledge	-.026	.015	3.02	.082
Age	.002	.0009	7.81	.005
(Scale)	.53			
Goodness of fit statistics^a	QIC=57.57		QICC=55.95	

Table 3.7. GEE results for model 3 with explicit PR rime as the outcome variable and implicit PR rime, vocabulary, letter-sound knowledge and age as predictors

Parameter	B	S.E.	Wald Chi-Square	Sig.
(Intercept)	-4.00	.70	32.36	<.0005
Implicit PR Rime	.39	.052	54.19	<.0005
Vocabulary	.21	.062	11.23	.001
Letter-sound knowledge	.002	.010	0.034	.85
Age	.003	.0005	22.71	<.0005
(Scale)	.27			
Goodness of fit statistics^a	QIC=32.57		QICC=33.52	

^a Computed using the full log quasi-likelihood function, in ‘small-is-better’ form

Table 3.8. GEE results for model 4 with explicit PR phoneme as the outcome variable and implicit PR phoneme, vocabulary, letter-sound knowledge and age as predictors

Parameter	B	S.E.	Wald Chi-Square	Sig.
(Intercept)	-2.92	.96	9.27	.002
Implicit PR Phoneme	.22	.068	10.23	.001
Vocabulary	.054	.078	0.48	.49
Letter-sound knowledge	.036	.011	10.51	.001
Age	.002	.0007	4.82	.028
(Scale)	.27			
Goodness of fit statistics^a	QIC=36.11		QICC=33.58	

^a Computed using the full log quasi-likelihood function, in ‘small-is-better’ form

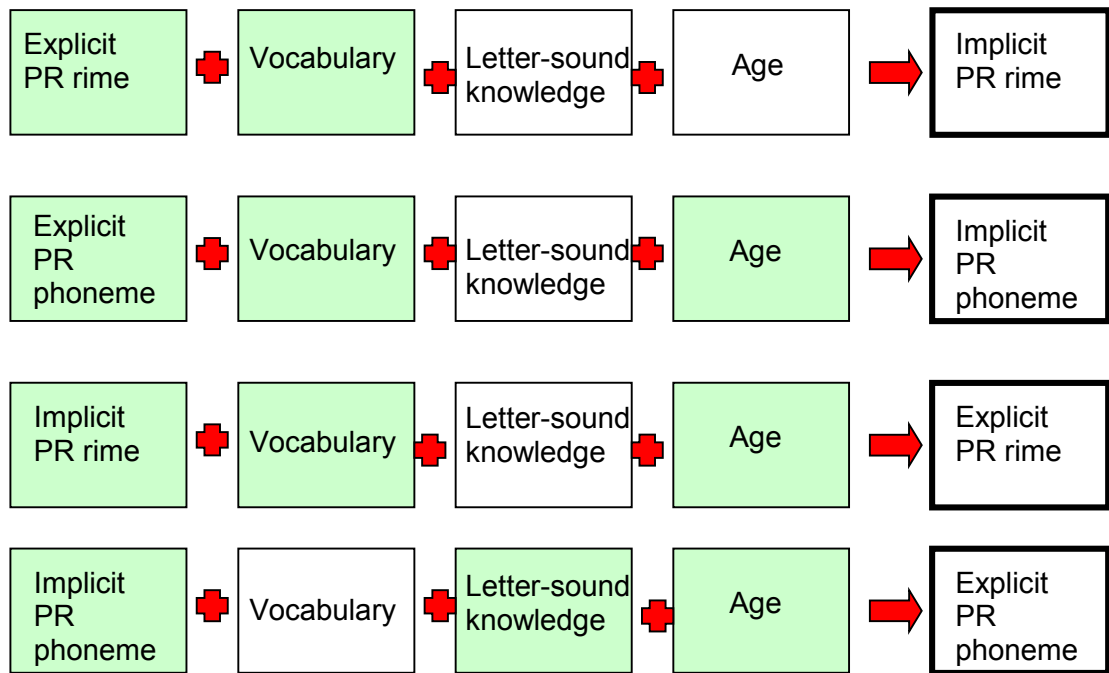


Figure 3.7. Summary of the GEE results. Boxes to the left of the arrows are predictors, while the boxes on the right are the outcome variables. Shaded boxes indicate significant predictors ($p < .05$)

Discussion

The current study aimed to explore the developmental trajectory of children's phonological knowledge during the first two years of school. Specifically, the study set out to test predictions made by four theoretical accounts (Table 3.1) through examination of the relative influence of vocabulary size and letter-sound knowledge on different aspects of phonological development. Concurrent use of implicit PR measures which do not require any explicit knowledge of phonological segments alongside explicit PR measures which directly question children on their knowledge of sound structure have allowed us to contrast the development of the segmentedness of the representations themselves versus the development of conscious phonological awareness. The use of a longitudinal design has provided the additional power needed to analyse children's performance on items designed to measure phonological representation at the rime and phoneme level separately, so that we can evaluate the four theoretical accounts more fully than has been previously possible.

The development of implicit and explicit knowledge of phonological representation

The study measured six key variables: implicit knowledge of phonological representation at both the rime and phoneme levels, explicit awareness of phonological representation (again at both the rime and phoneme levels), vocabulary size and letter-sound knowledge. All six variables were found to increase developmentally as we would expect, with a high level of interrelatedness found between them even when controlling for age. This confirms that there are important relations between these variables over and above general developmental factors, and is consistent with the previous cross-sectional study (Chapter 2). The only exception is the non-significant partial correlation (when controlling for age) between implicit PR at the phoneme level and letter-sound knowledge which supports Ventura and colleagues' proposition (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007) that segmental representations develop independent of literacy. The charts for the individual implicit PR measures (Figure 3.2) reveal a cleaner developmental trajectory than was found in the earlier cross-sectional study with all measures showing a steady increase over the first two years of school (although the trend did not reach significance for the pseudoword similarity task). This is likely to be due to the removal of between-subjects differences afforded by a repeated measure design and the elimination of the problem of overlapping age bands within the earlier study.

A key aim of the current study was to examine rime versus phoneme level performance separately on measures of phonological representation, motivated by the prediction of different causal pathways for rime and phoneme units within the three

variants of the emergent view (Table 3.1, see also Figure 3.1). Phoneme level performance on both implicit and explicit PR tasks lagged significantly behind rime level performance (with the exception of the pseudoword similarity task) which is consistent with the idea of an intermediate rime level of representation, as predicted by the lexical restructuring model (Metsala & Walley, 1998). Alternatively children may master rime level items earlier simply because of the natural salience of rimes relative to bodies within the English language (e.g. Treiman, 1992). In order to establish whether this finding does indeed reflect a shift in representational grain size, further work is needed using similar PR measures where frequency and neighbourhood density have also been manipulated. If children were found to make more segmental responses on items involving high frequency words from dense neighbourhoods (words which the lexical restructuring model proposes to be restructured first) than on items involving low frequency words from sparse neighbourhoods this would provide support for the idea that the observed lag in sensitivity to phonemes relative to rimes reflects a shift in representational grain size rather than just the natural salience of rhyme within English.

The role of vocabulary growth and letter-sound knowledge

The four theoretical accounts illustrated within Figure 3.1 differ in terms of the predicted roles of vocabulary growth and letter-sound knowledge on particular aspects of phonological development (Table 3.1). The accessibility view is the simplest account which proposes that children's phonological representations are fine grained from the outset but that children are only able to access phonemic segments at a conscious level once they have mastered the alphabetic principle (i.e. learnt the mappings between letters and sounds). Within the accessibility position vocabulary growth is not afforded any special role (as phonological representations are segmented from infancy), and letter-sound knowledge is proposed to be important for 'unlocking' conscious access to phonemes.

All three emergent accounts agree that rime emerges naturally as a linguistic unit at both a representational and conscious level through oral language experience alone. Where they differ is in their predictions about the emergence of the phoneme. While psycholinguistic grain size theory suggests that children need to be taught explicitly about the phonemes within words before they begin to organise their lexicon phonemically (Ziegler & Goswami, 2005), the LRM (Metsala & Walley, 1998) proposes that phonemes need to emerge at a representational level first driven by vocabulary growth, setting the stage for the development of later phoneme awareness. Ventura and colleagues extended the LRM by proposing that restructuring takes place in the absence of literacy but that

explicit awareness of phonemes only develops once the mappings between letters and sounds have been learnt (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007).

The GEE results obtained within the current study have allowed us to test the predictions summarised above by examining the relative roles of vocabulary growth and letter-sound knowledge on the four aspects of phonological development identified within Figure 3.1 (implicit PR rime/phoneme and explicit PR rime/phoneme). The results (summarised in Figure 3.7) show that vocabulary size is more predictive than letter-sound knowledge of performance on implicit measures of PR (at both the rime and phoneme level) and explicit measures of PR at the rime level. In contrast, letter-sound knowledge is more predictive than vocabulary of performance on explicit PR performance at the phoneme level. Given that the implicit PR measures are designed to probe PR segmentedness at the rime and phoneme level and that the explicit PR measures probe children's conscious rime and phoneme awareness, we can therefore infer that vocabulary size plays a more important role than letter sound knowledge in the segmentation of children's representations at both the rime and phoneme level and in the development of explicit rhyme awareness. Conversely, letter-sound knowledge seems to be more important than vocabulary growth for explicit phoneme awareness. The results are therefore most consistent with Ventura and colleagues (2007) proposition that phonemes may emerge naturally at a representational level, independent of orthographic knowledge, but that children only gain conscious access to phonemes once they have learnt a substantial number of letters. It is worth noting that the results are also consistent with the lexical restructuring model although this is underspecified relative to Ventura's account: while the lexical restructuring model proposes that vocabulary driven restructuring sets the stage for the emergence of phoneme awareness, no prediction is made as to whether letter-sound knowledge also play a role.

While much of the above discussion is concerned with how and when the phoneme emerges, it is important to note that there is still considerable debate about whether or not the phoneme actually exists –in the sense of a real psychological entity (e.g. Lotto & Holt, 2000). Phonemes are clearly a useful abstraction allowing the categorisation of sounds in terms of their ability to contrast meaning. However they are not separable within the acoustic speech stream and there is some evidence that words are stored in terms of subphonemic information (Lotto & Holt, 2000). Nevertheless, the phoneme remains the fundamental unit within the teaching of reading and PA skills within schools and so until the debate around what units should be considered fundamental to speech perception is resolved the study of children's sensitivity to phonemes remains a valid and important topic for research.

Conclusion

In conclusion, we have tested predictions made by the key theoretical accounts of phonological development using implicit and explicit PR measures designed to probe PR segmentedness and phonological awareness respectively. The results support the idea of developmental restructuring of children's phonological representations (e.g. Metsala & Walley, 1998) and are consistent with Ventura's proposition (Ventura et al, 2007) that while restructuring occurs in the absence of literacy, letter-sound knowledge is needed for the emergence of conscious phoneme awareness. The study provides further evidence of a dissociation between performance on implicit and explicit measures of phonological representation in terms of their ability to detect segmental sensitivity independent of orthographic knowledge. The study extends previous work by breaking down this dissociation into performance on rime versus phoneme level items and revealing that conscious awareness of phonemes but not rimes is dependent on letter-sound knowledge.

**GLOBAL VERSUS PHONEMIC SIMILARITY: EVIDENCE
IN SUPPORT OF MULTI-LEVEL REPRESENTATION**

Stephanie AINSWORTH

Stephen WELBOURNE

Anna WOOLLAMS

Anne HESKETH

Paper prepared for publication (not yet submitted)

Abstract

There is long standing debate about the extent to which children represent words in terms of global properties or phonological segments. Yet few studies have investigated how children's sensitivity to phonemic versus global similarity changes over time. The current study uses a mispronunciation reconstruction task to measure both types of sensitivity alongside one another within a cross-sectional (N=90, age range: 3;2 to 5;7) and longitudinal sample (N=24, overall age range 3;2 to 5;1). Results show that children's sensitivity to phonemes increases over the first two years of school but does not reach adult levels. Children's sensitivity to global similarity increases during the nursery year, reaching adult levels by the beginning of reception. The findings indicate that global similarity relations remain important throughout development and support the idea of multi-level representation.

Introduction

Global similarity – the degree to which words sound alike overall – has been highlighted as an important aspect of speech perception (e.g. Carroll & Snowling, 2001, Treiman & Breaux, 1982, Byrne and Fielding-Barnsley, 1993). A common theme within theoretical accounts of phonological development is the extent to which children represent phonological forms segmentally (i.e. in terms of phonological segments such as phonemes or onsets and rimes) or globally, in terms of indivisible wholes (Metsala & Walley, 1998; Walley, Metsala & Garlock, 2003; see Ainsworth, Welbourne, Woollams and Hesketh, submitted for an overview). The present study investigates developmental changes in the influence of global versus segmental similarity on children's similarity judgements and discusses the implications for theories of phonological development.

Global versus segmental representation

Within the lexical restructuring model it is argued that children's phonological representations (the way that children store the sound structure of words) are initially holistic: based on the overall 'sounds-likeness' of words or on one particularly salient feature (Metsala & Walley, 1998). As children learn more words their representations are proposed to become increasingly segmental (moving from whole word to onset-rime to phonemic representations) in order to keep phonologically similar words distinct. Evidence in support of a developmental shift from global to segmental representation comes from similarity judgement studies (Treiman & Baron, 1981; Treiman and Breaux, 1982) where children and adults heard three syllables and were asked to choose which two were most alike. While adults generally chose the pair of syllables sharing a common phoneme, children tended to choose the pair which shared no phonemes but were close to one another in terms of global similarity. Treiman and Breaux's findings (1982) have been interpreted as a reflection of qualitative differences between the phonological representations of adults and children (Carroll & Snowling, 2001), however it is worth noting that although adults displayed a bias towards phoneme versus global similarity relations, the bias was only small with mean proportions of .48 and .43 for the number of phonemic versus globally matched pairs chosen (the remaining .09 consisted of anomalous choices where syllables were not matched in terms of a common phoneme or global similarity). Treiman and Breaux's (1982) study therefore indicates that while adults are more sensitive to the number of shared phonemes than children are, their perception of how alike two syllables sound remains heavily affected by global similarity. A potential limitation of Treiman and Breaux's study (1982) is that the working memory demands of the task are relatively high with children (aged 3;6 to 5;5) required to hold all three

syllables in memory while mentally comparing each of the three possible pairs in terms of similarity. It is therefore unclear whether the lower ratio of phoneme to global responses found in children is due to children being more sensitive to the global properties of words than phonemes or is due at least in part to the fact that children struggled with the working memory demands of the task.

Further support for a whole to parts shift in phonological representation comes from developmental differences in position effects on mispronunciation detection tasks. While adults tend to identify mispronunciations more easily when they are at the beginning of the word, children are found to be less affected by the position of the mispronunciation (Cole & Perfetti, 1980; Walley, 1987). Bowey & Hirakis (2006) however found that such position effects covary with acoustic clarity and that once this confound is controlled for, position effects are found in both adults and children. The idea of children's representations being initially based on global properties and then becoming increasingly componential fits in well with the shift from holistic to dimensional perception of objects found within the visual domain (Shepp & Seartz, 1976, Smith & Kemler, 1977) and is also consistent with observational studies of language acquisition (Ferguson & Farwell, 1975). However few studies have tested the hypothesis directly and, as discussed above, they have been limited in terms of effect size and/or methodological issues. There is therefore a need for tasks which measure sensitivity to phonemic versus global similarity where working memory task demands are minimised.

PRIMIR (Processing Rich Information from Multidimensional Interactive Representations) provides an alternative account of phonological development where global similarity remains an important influence on the perception of phonological forms even after children have become sensitive to phonological segments (Werker & Curtin, 2005). Within PRIMIR, words are simultaneously represented across three multidimensional planes: perceptual, word form and phonemic. While the perceptual plane stores all information contained within the acoustic signal, the word form plane segments words from the speech stream and stores the phonetic and indexical information within word-level exemplars (Werker & Curtin, 2005). As infants' vocabularies grow, higher order regularities begin to emerge and phonemic categories form within the phonemic plane. Initially these categories will not map directly onto adult phonemic categories: phonemic categories will become mature only after they have been 'sharpened up' by the process of learning to read (Werker & Curtin, 2005). Within PRIMIR the extent to which different levels of information are processed during a particular task depends on three dynamic filters: the child's initial biases, the child's developmental level and the demands of the linguistic task. These three filters act as a lens with different aspects of the rich

information available to the child coming in and out of focus over development and for different tasks.

The PRIMIR framework was motivated by a need to account for conflicting findings within the speech perception literature about the level of detail within infants' representations of words. There are numerous examples within the infant literature where performance on one task suggests representation at one level, while performance on another task indicates representation at a different level (Werker & Curtin, 2005). For example, on counting tasks infants use syllables, whereas on discrimination (Dupoux & Peperkamp, 2002; Miller & Eimas, 1996), preference (Jusczyk, Goodman & Baumann, 1999) and segmentation tasks (Christophe, Dupoux, Bertoncini & Mehler, 1994; Jusczyk, Hohne & Bauman, 1999) infants of the same age use information stored at the subsyllabic level. PRIMIR accounts for these findings by proposing that representations exist at multiple levels throughout development with some levels of information being more salient than others depending on the level of development of the child and the task at hand (Werker & Curtin, 2005).

Although Werker and Curtin (2005) do not refer to the idea of global similarity directly, it can be framed in terms of the clustering of word level exemplars within the word form plane. Within PRIMIR the gradual emergence of phonemes does not result from the restructuring of holistic representations into a more segmental form (as it does in the lexical restructuring model (Metsala & Walley, 1998)) – rather the word form plane remains throughout development with phonemes extracted from regularities found within it and stored within a separate representational space. Within the PRIMIR framework we might then expect global similarity to continue to have an important influence on speech perception throughout development, alongside an emerging sensitivity to phonemes. Conversely within the lexical restructuring model, we might expect that as representations become increasingly segmental children's sensitivity to phonemes will rise at the expense of a decrease in sensitivity to global similarity. The two hypotheses have yet to be tested empirically.

Developmental changes in global sensitivity

The concept of global similarity is central to theoretical accounts of phonological development. Yet little work has been done to investigate if and how children's sensitivity to the global properties of words changes over time. The majority of work has focussed instead on the emergence of segmental sensitivity (sensitivity to phonemes and rimes (e.g. Ainsworth, Welbourne, Woollams & Hesketh, submitted b; Foy & Mann, 2009; Storkel, 2002). Where researchers have attempted to explore children's sensitivity to global

similarity there have been issues with task design. For instance within Wagensveld's rhyme judgement task (Wagensveld, Segers, Alphen, van & Verhoeven, 2013) the influence of global similarity was confounded with phonemic similarity. When asked to identify whether or not a set of words rhymed with a given target (e.g. 'bek'), both the rhyming choice ('gek') and non-rhyming but globally similar choice ('bak') shared two phonemes with the target. Wagensveld et al. (2013) interpreted the finding that adults chose the non-rhyming globally matched items more often than children did as evidence that global sensitivity increases over time. However, it could be that adults chose 'bak' more often than children did because of an increased sensitivity to phonemes. The current study manipulates both global similarity and the number of shared phonemes to allow us to separate the two types of sensitivity and plot them alongside one another over the first two years of school (age range 3;2 to 5;7). Although no direct predictions are made within the literature about how global sensitivity might change over time, three logical possibilities are considered here:

1) **Children's sensitivity to global similarity decreases as they get older.**

This possibility would be consistent with the qualitative change in the structure of children's phonological representations predicted by the lexical restructuring model (Metsala & Walley, 1998). If children's phonological representations are initially holistic and become increasingly segmental over development we might expect that as children's segmental sensitivity rises we would see a corresponding drop in global sensitivity.

2) **Children's sensitivity to global similarity increases as they get older.**

This possibility fits in well with PRIMIR's idea of word level exemplar-based representation remaining important for speech perception even after the emergence of segmental representation within the phonemic plane (Werker & Curtin, 2005). Because the PRIMIR framework allows for simultaneous representation at different levels (rather than replacement of global representation by segmental representation) we would not expect a corresponding decrease in levels of global sensitivity. We would also expect children's sensitivity to global similarity to increase over development as more exemplars are added to the word form plane.

3) **Children's sensitivity remains stable.** Although this pattern is not predicted by any current theory of phonological development it is included here as a logical possibility given the lack of previous work on the development of sensitivity to global similarity.

The current study aims to test which of the three developmental trajectories listed above best reflects children's similarity judgements on a mispronunciation reconstruction task. The task includes four response choices which have been manipulated in terms of both global and phonemic similarity relations allowing us to separate these two potential influences on performance. This is the first time (to the authors' knowledge) that both phoneme and global sensitivity have been investigated concurrently within the same sample.

Method

Participants

90 children were recruited from two schools of low and medium socio-demographic status respectively (as indexed by the percentage of children receiving free school meals, see Ofsted, 2012). Children were included within the study if they had at least one English speaking parent living at home with them and if they had no known hearing problems or special educational needs (as reported by their class teacher). All children spoke fluent English. We do not have complete information about the children's home language environments, although for the large majority of children within the study English was the child's first language. All children were tested as part of a larger scale cross-sectional study of phonological development (Chapter 2). The youngest 24 children were tested as part of a longitudinal study over three additional time points roughly 5 months apart (Chapter 3). 74 adult undergraduate students were also included as a control group. The adults were tested as a part of a wider battery of tests (see Chapter 2). Adult participants were excluded if English was not their first language or if they had any diagnosed dyslexia or hearing difficulties.

Procedure

Children were tested in a quiet room within school. Corrective feedback was given for one training item at the beginning of the task. For all other items general praise and stickers were given as encouragement independent of performance.

Materials

Items were selected as being familiar to young children – 33 of the 36 words used can be found within Storkel and Hoover's database (Storkel & Hoover, 2010) which is based on corpora of kindergarten and first grade children (words not found were wheel, nose and horse). For every trial children were first asked to name the pictures and were told the name if they were unable to identify the picture.

All the distracters within the task were matched listwise in terms of frequency and phonotactic probability. Two phonotactic probability statistics were used: positional segment average (how often the segments within a word occur in that position within other words) and biphone average (the frequency of pairs of sound segments) using Storkel and Hoover's online calculator (Storkel & Hoover, 2010). See Appendix 2 for the stimuli list and Appendices 8 to 10 for additional details on matching.

The mispronunciation reconstruction task

Children heard a puppet mispronounce a CVC word and were then asked to guess which picture he was trying to say – which picture did it sound the most like. For example the puppet said ‘*hain*’ and the children chose whether he was trying to say rain, pin, bone or tap. For each trial the child was presented with four response choices:

- 1) Two-phoneme response: a word sharing two phonemes with the stimulus (e.g. *nake-name*)
- 2) One-phoneme globally matched response: a word sharing only one phoneme with the stimulus but matched with the two-phoneme response in terms of global similarity to the stimulus (e.g. *nake-net*).
- 3) One-phoneme unmatched response: a word sharing one phoneme with the stimulus and of lower global similarity to the stimulus than choices 1) and 2) (e.g. *nake-nurse*).
- 4) Unrelated response: this word shares no phonemes with the stimulus and is also globally distant (e.g. *nake-shed*).

The rationale behind the task is that if children choose the closest phonemic match – i.e. the two-phoneme response – more often than the one-phoneme globally matched response, then we can infer that they are sensitive to the number of shared phonemes over and above how close the words are in terms of global similarity. Similarly if children choose the one-phoneme globally matched response more often than the one-phoneme unmatched response we can infer that they are sensitive to global similarity when phonemic similarity is held constant.

Global similarity matching

For each trial the two-phoneme response and the one-phoneme globally matched response were matched in terms of global similarity. For example when children are asked to choose whether ‘rain’ or ‘pin’ sounds the most like ‘*hain*’, ‘pin’ is just as close to ‘*hain*’ in terms of global similarity despite sharing only one (rather than two) phonemes with ‘*hain*’. Global similarity scores were calculated using adult ratings collected by Singh and colleagues (Singh & Woods, 1971; Singh, Woods & Becker, 1972). In the first study (Singh & Woods, 1971) adults were asked to rate how similar pairs of vowels were when presented in isolation (e.g. *i* and *ε*) on a scale of 1 to 7. In the second study (Singh, Woods & Becker, 1972) adults were asked to rate the similarity of pairs of consonants when presented in front of the vowel /*ɑ*/ (e.g. how similar are /*ba*/ and /*pɑ*/) again on a scale of 1

to 7. Within the current study, scores of how dissimilar stimuli were from one another were calculated from the Singh data using the same method adopted by Treiman and Breaux (1982) and Carroll and Snowling (2001). For example, the dissimilarity score between the words 'pin' (pronounced /pɪn/) and 'bed' (pronounced /bɛd/) is the dissimilarity of /p/ and /b/ (3.9) plus the dissimilarity of /ɪ/ and /ɛ/ (2.22) plus the dissimilarity of /n/ and /d/ (4.8).

Results

How do response patterns change over development?

Plots of response frequency by response type are shown in Figures 4.1 and 4.2 for the longitudinal and cross-sectional data sets respectively. Both figures show a steady developmental increase in the number of two-phoneme responses. A repeated measures ANOVA conducted on the longitudinal data revealed a significant main effect of age on the number of two-phoneme responses, $F(3,66)=16.58$, $p<.0005$. Repeated within-subjects contrasts showed a significant increase from time point 1 to time point 2 ($F(1,22)=5.19$, $p=.033$), and a significant increase from time point 2 to time point 3 ($F(1,22)=7.14$, $p=.014$). The increase in two-phoneme responses from time point 3 to time point 4 failed to reach significance, $F(1,22)=3.41$, $p=.078$. A one way ANOVA conducted on the cross-sectional data also found a significant main effect of age on the number of two-phoneme responses, $F(3,86)=4.04$, $p=.01$. Pairwise comparisons (with Bonferroni correction) between adjacent age groups were all non-significant: young nursery to old nursery, $t(22)=.96$, $p=.88$, old nursery to young reception, $t(22)=.79$, $p=1.00$, young reception to old reception, $t(22)=.45$, $p=1.00$. In both cases and for all subsequent analyses, unless otherwise stated, only the child data (and not the adult data) were included.

The rise in two-phoneme responses is accompanied by a corresponding decrease in the unrelated response category as we might expect (longitudinal: $F(2,12, 46.69)=5.13$, $p=.009$, cross-sectional: $F(3,86)=5.00$, $p=.003$). Repeated within-subjects contrasts conducted on the longitudinal anomalous responses were however all non-significant (time point 1 to time point 2: $F(1,22)=4.16$, $p=.054$; time point 2 to time point 3: $F(1,22)=1.06$, $p=.31$; time point 3 to time point 4: $F(1,22)=0.46$, $p=.50$), Pairwise comparisons (with Bonferroni correction) conducted on the cross-sectional anomalous responses were also all found to be non-significant (young nursery to old nursery: $t(86)=.33$, $p=1.00$; old nursery to young reception: $t(86)=0.48$, $p=1.00$; young reception to old reception: $t(86)=0.57$, $p=.86$).

Figures 4.1 and 4.2 also show a decreasing trend in the frequency of one-phoneme unmatched responses as confirmed by a significant main effect of time in the longitudinal ($F(3,66)=6.02$, $p=.001$) but not the cross-sectional data ($F(3,86)=1.98$, $p=.12$). Repeated within subjects contrasts conducted on the longitudinal one-phoneme unmatched responses were all non-significant (time point 1 to time point 2: $F(1,22)=2.48$, $p=.13$; time point 2 to time point 3: $F(1,22)=2.55$, $p=.124$; time point 3 to time point 4: $F(1,22)=.19$, $p=.67$). The number of globally matched responses remained relatively stable with no significant effect of time found on the number of one-phoneme globally matched responses for either the

longitudinal ($F(2.35,51.7)=2.20, p=.11$) with Greenhouse-Geisser correction) or cross-sectional data ($F(3,86)<1, p=1.00$).

As expected adults chose the two-phoneme response most of the time ($M=74.65\%$, $S=13.81\%$). Their response pattern suggests that adults remain affected by global similarity, choosing the one-phoneme globally matched response significantly more often than the one-phoneme unmatched response, $t(71)=3.21, p=.002$.

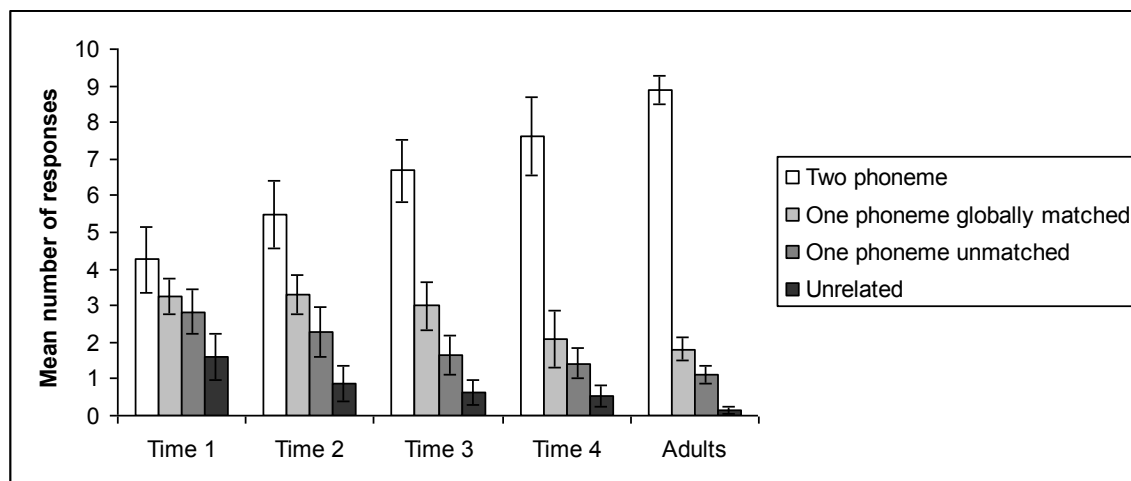


Figure 4.1. Children’s response profiles across four different time points with adult performance included for comparison. Error bars represent confidence intervals of 95%.

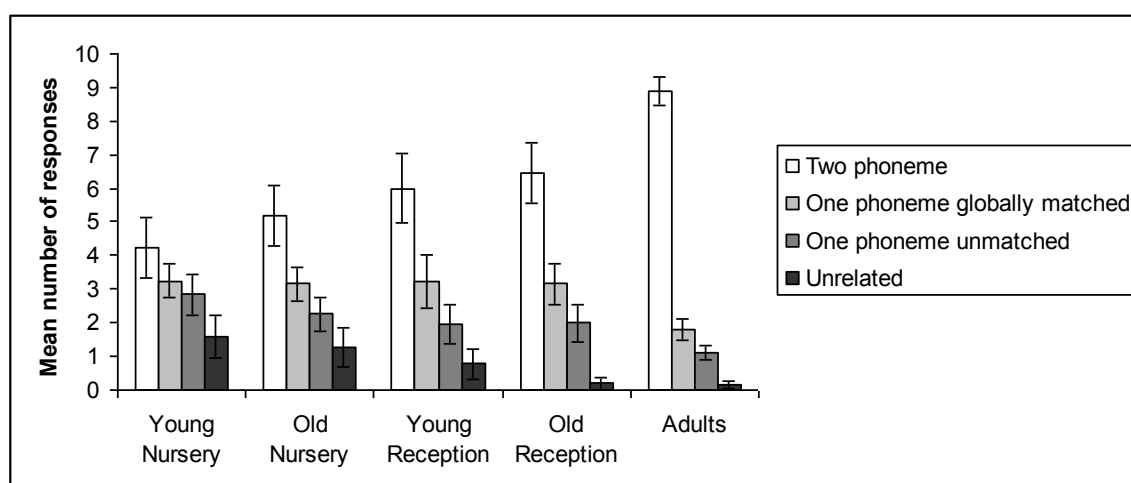


Figure 4.2. Children’s response profiles across four cross-sectional groups with adult performance included for comparison. Error bars represent confidence intervals of 95%.

Sensitivity to phonemic versus global similarity

Taken at face value, the above pattern of results might be interpreted as evidence of a developmental rise in sensitivity to phonemes alongside static levels of global sensitivity. However, the frequency of a particular response type does not tell us directly about sensitivity unless we take into account the effect of the other available response types. In order to measure how sensitive children are to the number of shared phonemes we need to look at how often they choose the word which shares the most phonemes while holding global similarity constant. In other words we need to compare how often children choose the two-phoneme response versus the one-phoneme globally matched response. To this end a measure of phoneme sensitivity was calculated according to the following formula:

(1) *Phoneme sensitivity*

$$= \frac{\text{no. of 2 phoneme responses}}{\text{no. of 2 phoneme responses} + \text{no. of 1 phoneme globally matched responses}}$$

Similarly, to assess how sensitive children are to global similarity we need to look at how often children choose the closest response in terms of global similarity while holding the number of shared phonemes constant. To do this we calculated the proportion of one-phoneme globally matched responses relative to one-phoneme unmatched responses as shown below:

(2) *Global sensitivity*

$$= \frac{\text{no. of 1 phoneme responses}}{\text{no. of 1 phon. globally matched responses} + \text{no. of 1 phon. unmatched responses}}$$

Phoneme and global sensitivity scores were calculated for each time point within the longitudinal study (Figure 4.3) and for each age group within the cross-sectional study (Figure 4.4). Analyses were performed on group rather than individual data to avoid the issue of zero frequencies yielding division by zero errors (where both frequencies within the denominator are zero) and skewing sensitivity scores (when the numerator only is zero). For each age group/time point the total number of responses within a given type summed over all participants within the group was used to calculate the sensitivity scores.

Binomial tests were conducted on the phoneme sensitivity scores for the youngest group of children (this corresponds to the young nursery children in the cross-sectional study who also took part in the longitudinal study) to see if they were sensitive to the

number of shared phonemes over and above global similarity. Similarly binomial tests were conducted on the youngest children's global sensitivity scores to see if they are sensitive to global similarity over and above the number of shared phonemes. While the youngest group's sensitivity to the number of shared phonemes was found to be significantly above chance ($p=.031$), their sensitivity to global similarity was found not to be significantly above chance ($p=.18$). However when the cross-sectional children were treated as one cohort (i.e. all age groups combined) children's sensitivity to both phonemes ($p<.0001$) and global similarity ($p=.00017$) were both found to be significantly above chance ($p<.0001$). As expected the adults' phoneme sensitivity level was well above chance ($p<.0005$). Adults were also found to have above chance levels of global sensitivity ($p<.0005$).

How does sensitivity to phonemic and global similarity change over development?

The longitudinal data (Figure 4.3) suggest developmental increases in sensitivity to both phonemic and global similarity. To establish the significance of these trends, sign test analyses were conducted on the longitudinal data at the item level. Each response for each item was first coded according to response type (i.e. 1=two-phoneme response, 2=one-phoneme globally matched, 3=one-phoneme unmatched, 4=unrelated). The responses were then filtered to allow us to isolate sensitivity to phoneme and global similarity respectively. In the former case we were interested in the contrast between two-phoneme responses and one-phoneme globally matched responses where the number of shared phonemes differs but global similarity is held constant. A filter was therefore applied for each pair of time points to allow analysis of only those cases where the response at the later time point was either a two-phoneme or one-phoneme globally matched response. For each item answered by each participant a phoneme sensitivity change score (of -1 0 or +1) was then given depending on whether the later response was less similar phonemically, equally similar phonemically or more similar phonemically than the previous response (see Table 4.1). Sign tests conducted on the phoneme sensitivity change scores for each consecutive time point showed a significant rise in phoneme sensitivity from time point 1 to time point 2 ($z=6.78$, $p<.0005$, one-tailed), time point 2 to time point 3 ($z=6.03$, $p<.0005$, one-tailed) and time point 3 to time point 4 ($z=5.36$, $z<.0005$, one-tailed).

Table 4.1. Phoneme and global sensitivity change scores assigned to longitudinal response patterns

Phoneme sensitivity change			Global sensitivity change		
Earlier response	Later response	Change coding	Earlier response	Later response	Change coding
two-phoneme	one-phoneme	less similar,	one-phoneme globally	one-phoneme	less similar
	globally matched	phonemically, -1	matched	unmatched	globally, -1
one-phoneme	one-phoneme	equally similar	two-phoneme	one-phoneme	less similar
globally matched	globally matched	phonemically, 0		unmatched	globally, -1
one-phoneme	one-phoneme	equally similar	one-phoneme globally	one-phoneme globally	equally similar
unmatched	globally matched	phonemically, 0	matched	matched	globally, 0
two-phoneme	two-phoneme	equally similar	one-phoneme unmatched	one-phoneme	equally similar
		phonemically, 0		unmatched	globally, 0
unrelated	one-phoneme	more similar	two-phoneme	one-phoneme globally	equally similar
	globally matched	phonemically, +1		matched	globally, 0
unrelated	two-phoneme	more similar	one-phoneme unmatched	one-phoneme globally	more similar
		phonemically, +1		matched	globally, +1
one-phoneme	two-phoneme	more similar	unrelated	one-phoneme globally	more similar
globally matched		phonemically, +1		matched	globally, +1
one-phoneme	two-phoneme	more similar	unrelated	one-phoneme globally	more similar
unmatched		phonemically, +1		unmatched	globally, +1

A similar set of sign tests was conducted to measure the developmental change in sensitivity to global similarity. This time, responses were filtered to include only those cases where the later response for a given pair of time points was either a one-phoneme globally matched response or a one-phoneme globally unmatched response. This allowed us to examine children’s sensitivity to global similarity while holding the number of phonemes constant. The filtered responses were then assigned a global sensitivity change score from one time point to another depending on whether the later response was less globally similar, equally globally similar or more globally similar (see Table 4.1). A series of sign tests conducted on the global sensitivity change scores revealed that there was a significant increase in global sensitivity from time point 1 to time point 2 ($z=1.71$, $p=.044$, one-tailed) however both the increase from time point 2 to time point 3 ($z=0.58$, $p=.28$, one-tailed) and the apparent dip from time point 3 to time point 4 ($z=1.01$, $p=.31$, one-tailed) were not significant.

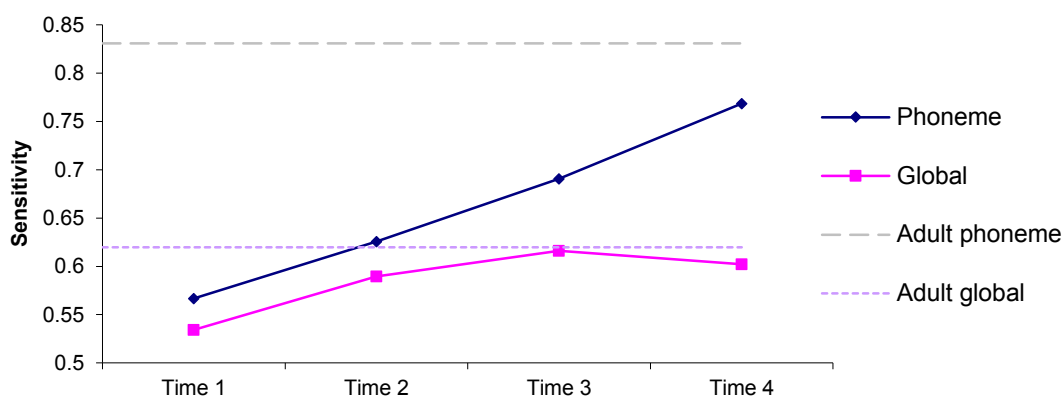


Figure 4.3. Developmental plot of children’s sensitivity to phonemes and global similarity from the longitudinal data. Adult sensitivity levels are included as target lines for comparison

The sensitivity plot for the cross-sectional data (Figure 4.4) shows a similar trajectory with an apparent rise in sensitivity to both phonemic and global similarity over development. To test the significance of the observed increases we conducted a series of binomial tests on the cross-sectional data (these analyses would have been inappropriate for use with longitudinal data given the lack of independence between responses at different time points). Phoneme sensitivity was found to increase beyond nursery, i.e. when comparing the phoneme sensitivity scores of the youngest children with the children’s responses combined across the three older groups ($p<.0005$). The increases between older nursery and reception children ($p=.16$) and between the young and older

reception groups ($p=.24$) were both found to be non-significant. Phoneme sensitivity was however found to increase between the older reception children and adults ($p<.0005$).

Binomial tests conducted on the global sensitivity scores within the cross-sectional data showed a significant rise beyond young nursery ($p=.00066$) and a marginally significant increase beyond older nursery ($p=.061$). There was no significant change between younger and older reception children ($p=.56$) and between older reception children and adults ($p=.38$). While young nursery children's global sensitivity was significantly lower than adults ($p=.0049$), the older nursery children's global sensitivity was only marginally below adult levels ($p=.055$).

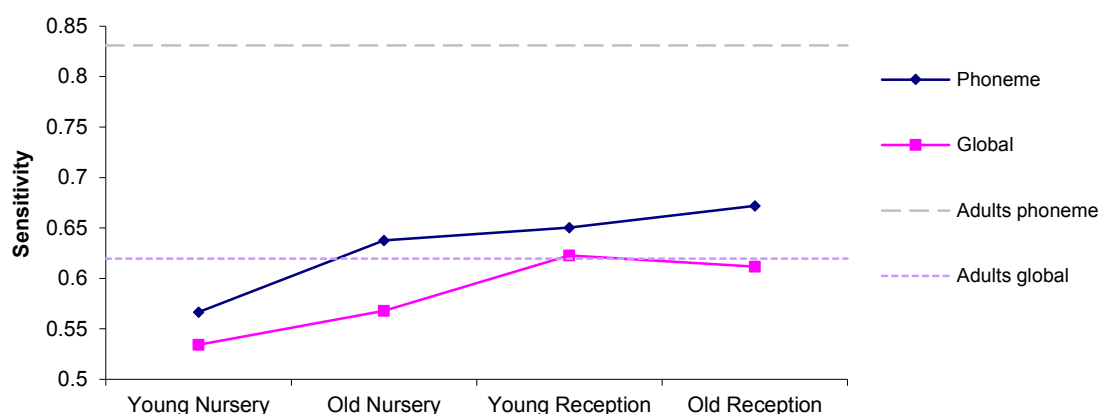


Figure 4.4. Developmental plot of children's sensitivity to phonemes and global similarity from the cross-sectional data. Adult sensitivity is included as target lines as a comparison

In summary, the results provide evidence that young nursery children (aged 3;2 to 3;10) are sensitive to the number of shared phonemes over and above global similarity. They are also sensitive to changes in global similarity when the number of shared phonemes is held constant. The longitudinal and cross-sectional data provide convergent evidence of a developmental increase in both types of sensitivity, although there are differences in terms of the exact trajectories. While both sets of data indicate an early increase in global sensitivity before levelling off around the same level found in adults, there is some disagreement in terms of the development of phoneme sensitivity. While the cross-sectional increase in phoneme sensitivity was only found to be significant early on in development, the longitudinal data indicates a steady increase in phoneme sensitivity throughout nursery and reception. Figures 4.3 and 4.4 have been plotted on the same scale to highlight the fact that the phoneme sensitivity levels of the early and late reception

groups are lower (4% and 10% lower respectively) than those reached by the longitudinal group at similar points within the reception year (time points 3 and 4). This suggests that cohort effects may underlie the discrepancy between the developmental trajectories for phoneme sensitivity within the cross-sectional and longitudinal data.

Discussion

The current study has contrasted the development of two types of sensitivity implicated within theories of phonological development: sensitivity to phonemes and sensitivity to global similarity. Sensitivity was measured using a mispronunciation reconstruction task where the response choices were manipulated in terms of both phonemic and global similarity, allowing us to separate the two potential influences on children's similarity judgements. Children aged 3;2 to 3;10 who had just started school already showed significant levels of sensitivity to the number of shared phonemes between words over and above how similar they were globally. Within the longitudinal data phoneme sensitivity increased steadily over time but was yet to reach adult levels by the end of reception.

Within the cross-sectional data, phoneme sensitivity was also found to increase, although the rise was only found to be significant when comparing the youngest age group (aged 3;2 to 3;10) with the oldest age group (aged 4;7 to 5;7). There are two main reasons why the increase in phoneme sensitivity might be less pronounced within the cross-sectional data. Firstly while the longitudinal data is collected from the same children at each time point, the cross-sectional data is collected from four different groups of children. The shallower gradient for the phoneme sensitivity plot within Figure 4.4 may therefore be due to cohort differences between the cross-sectional nursery children – whose phoneme sensitivity levels are in line with those of the longitudinal group at corresponding points within the school year (i.e. time points 1 and 2) - and the cross-sectional reception children who had lower levels of sensitivity than the reception children within the longitudinal group at time points 3 and 4. For example, the cross-sectional 'young reception' and 'old reception' groups might be less sensitive to phonemes than expected because of differences in teaching method used within their classrooms or because they are generally a less able cohort of children.

The second possible explanation for the discrepancy between the cross-sectional and longitudinal data is that the longitudinal group showed a greater effect of time on phoneme sensitivity because they benefitted from the study itself. While the cross-sectional participants only took part in each task once, the longitudinal participants took part in each task four times. It is therefore possible that this extra experience of listening to and reflecting on the sounds within words, could have produced a training effect in the longitudinal group. However, given the lack of an explicit training element and the fact that corrective feedback was given for one training item only, it is unlikely that participation in the study would be the main reason behind the steady developmental increase in phoneme sensitivity.

Given the limitations of cross-sectional data within a developmental context and the fact that training effects are judged as being unlikely it is therefore assumed that the longitudinal data provides a more reliable reflection of the development of phoneme sensitivity than the cross-sectional data. The observed rise in phoneme sensitivity over time is consistent with performance on other tasks which control for global similarity reported within Chapters 2 and 3. It is also consistent with the lexical restructuring model (Metsala & Walley, 1998), the PRIMIR framework (Werker and Curtin, 2005), and other emergent theories of phonological development (e.g. Ziegler & Goswami, 2005; Ventura, Kolinsky, Fernandes, Querido & Morais, 2007) which predict phonemic representation to emerge gradually over development.

Both the longitudinal and cross-sectional data also found global sensitivity to increase significantly over time, although the rise was limited to the early stage of the developmental period studied. By the second half of the nursery year global sensitivity levels flattened off, approximating adult levels. The fact that global sensitivity did not drop over development suggests that there is no ‘trade off’ associated with the rise in phoneme in sensitivity as we might expect within the lexical restructuring model (Metsala & Walley, 1998), where global representations are restructured into a more segmental form. The observed rise in both phoneme and global sensitivity is more consistent with PRIMIR’s idea of word level exemplar based representations (within the word form plane) remaining important as phonemic categories emerge within the phoneme plane (Werker & Curtin, 2005). The fact that children’s global sensitivity increases rather than remaining stable may be explained in terms of exemplars being added to the word form plane as children gain more language experience. The results suggest that although adults are much more sensitive to phonemes than children are, their classifications are still influenced by global similarity as evidenced by the fact that adults were not at ceiling on the mispronunciation reconstruction task, and made globally matched one-phoneme responses significantly more often than unmatched one-phoneme responses. Again, this is consistent with PRIMIR’s idea of simultaneous levels of representation throughout development.

It is important to note that the main findings of this paper rely on the assumption that we have been able to isolate phoneme and global sensitivity by manipulating the number of shared phonemes and the global similarity distance between words. While it is relatively trivial to manipulate phonemic similarity, the operationalization of global similarity is more contentious. As highlighted by Luce and Pisoni (1998) the idea of similarity in speech perception is poorly defined and may be conceptualised in a number of different ways (Kessler, 2005). Within the current paper and other similar studies within the phonological representation and phonological awareness literature (Byrne & Fielding-

Barnsley, 1993; Carroll & Snowling, 2001) the global similarity distance between words is calculated using adult judgement data collected by Singh and data (Singh & Woods, 1971; Singh, Woods, & Becker, 1972). This method of calculating global similarity has the benefit of being grounded empirically: rather than making assumptions about how adults make judgements about inter-phonemic similarity (cf. the use of theoretical feature based systems, e.g. Connolly, 1997; Bailey & Hahn, 2005; Harm and Seidenberg, 1999; Li and MacWhinney, 2002) we directly measure them. It also has the benefit of providing a straightforward, linear metric which allows items to be quickly and easily matched. There are however, a number of potential problems with this method both in terms of the validity of the ratings themselves and the way in which they are used to calculate global similarity distance.

Firstly, judgement based data is inherently subjective and may be very sensitive to the specific task and instructions. With regard to ratings based judgements in particular, Kessler (2005) points out that while people are good at making categorical judgements about words, they tend to be inconsistent at rating word similarity along a continuous or ordinal scale. Secondly, the data is limited in scope by the fact that Singh and colleagues tested adults on only 22 consonants and 12 vowels (Singh & Woods, 1971; Singh, Woods, & Becker, 1972). This meant that we were only able to construct matched items for the mispronunciation reconstruction task which contained combinations of these phonemes. The number of trials that we were able to include within the task was thus limited. Thirdly, and perhaps most importantly, the assumption that the similarity between two words can be calculated by treating words as strings of phonemes and adding together the scores on a phoneme by phoneme basis is problematic on both theoretical and empirical grounds.

Theoretically, one might argue that it is inappropriate when attempting to measure 'global' similarity to use a method which involves abstraction into discrete phonemes. The whole idea behind global similarity, proposed to be used in young children's categorisations of words, is that words are compared as indivisible wholes and that analysis is not undertaken on a phoneme by phoneme basis. When we say a word the phonemes are not spoken discretely one after the other – rather they are run together with a degree of overlap between them. This overlap, otherwise known as coarticulation, allows communication to be more efficient but involves the modification of phonemes from the form that they have when spoken in isolation (Liberman & Mattingly, 1989). Coarticulation is context dependent with different neighbouring sounds causing different modifications to the way that a phoneme is spoken. For example the /b/ in /bi/ is realized differently than in /ba/ - with the jaw being involved more in the closing of the lips for

close vowels relative to open vowels (Fowler & Galantucci, 2005). Similarly the /s/ in ‘spot’ sounds different to the /s/ in ‘scot’ because the labial consonant /p/ modifies /s/ in a slightly different way to the velar consonant /k/ (Stevens, 1998, p.558-561). We can therefore not assume that for example /b/ and /s/ will have the same degree of global similarity regardless of which words they are in. In order to overcome this limitation, and potentially match items more accurately we would need to have a set of values for every phoneme pair in every position. However to the author’s knowledge no such data set currently exists and would be very time consuming to collect.

While we acknowledge that the method used to operationalise global similarity within the present study has a number of limitations, we believe that the study remains a promising first step towards isolating two theoretically distinct influences – sensitivity to phonemes and sensitivity to global similarity – on phonological representation and speech perception more generally. The development of a more accurate measure of global similarity distance which takes into account the effects of coarticulation remains a useful avenue for future research.

Conclusion

In summary this study has used a mispronunciation reconstruction task to plot the development of two important influences on speech perception: phoneme sensitivity and global sensitivity. We have presented evidence that three-year-olds already have significant levels of both phoneme sensitivity and global sensitivity as they enter school. While children are just as sensitive to global similarity as adults are from the second half of nursery onwards (aged 4;0 to 4;5), they are yet to reach adult levels of phoneme sensitivity by the end of reception (aged 4;6 to 5;1). The results show a developmental rise in sensitivity to both phoneme and global similarity relations. These findings are consistent with PRIMIR’s idea of word level exemplar based representations being augmented throughout development alongside an emerging phoneme plane (Werker & Curtin, 2005).

**EVIDENCE OF EMERGING SEGMENTAL STRUCTURE
WITHIN A NEURAL NETWORK MODEL TRAINED USING
REAL SPEECH DATA**

Stephanie AINSWORTH

Stephen WELBOURNE

Anna WOOLLAMS

Anne HESKETH

Paper prepared for publication (not yet submitted)

Abstract

There is currently debate about the nature of children's phonological representations and the potential roles of literacy and vocabulary growth within their development. Neural network models allow us to simulate phonological development while directly controlling oral language experience and exposure to literacy. While most previous models have used artificial inputs and outputs which potentially bias the network towards phonemic representation, the current model is trained on real speech data allowing us to look for evidence of naturally emerging segmental representation. The feed forward model was successfully trained on the mappings between acoustic recordings of spoken words and the corresponding articulatory measurements (laryngograph, electromagnetic-articulograph and electro-palatograph). Analysis of the model's hidden unit representations was conducted using a novel 'hidden plot distance' metric. The distance between word representations became increasingly sensitive to both phonemic and global similarity as the model developed. The implications for current theories of phonological development are discussed.

Introduction

Current theories of phonological development make different predictions about the emergence of segmental phonological representation (e.g. Chapter 2; Walley, Metsala & Garlock, 2003; Ziegler & Goswami, 2005). In particular there is debate surrounding the extent to which preliterate children represent words in terms of global properties or phonological segments (e.g. see Chapter 4; Metsala & Walley, 1998). While proponents of the accessibility view propose that children's representations are adult-like from infancy (Rozin & Gleitman, 1977; Liberman & Shankweiler & Liberman, 1989; Bailey & Plunkett, 2002, Ballem & Plunkett, 2005, Swingley & Aslin, 2000; Swingley, 2009), emergent accounts propose that children's phonological representations do not become phonemic until later in childhood (Metsala & Walley, 1998; Ziegler & Goswami, 2005).

Within the lexical restructuring model it is proposed that children's phonological representations are initially holistic based on the overall acoustic form of the word or on one particularly salient feature (Metsala & Walley, 1998). The idea here is that when children's vocabularies are limited they do not need to store words in much detail in order to keep them distinct. As they learn more words however, children's representations need to become increasingly detailed and componential to avoid confusion between similar sounding words. Within this account representations are predicted to be restructured gradually with high frequency words from dense neighbourhoods (words with many phonologically close neighbours) restructured earlier than low frequency words from sparse neighbourhoods (words with few phonologically close neighbours).

Although the lexical restructuring model makes no direct predictions about the role of letter-sound knowledge (Metsala & Walley, 1998), the emphasis is on oral language experience as the key driver for representational change. Ventura and colleagues extended this idea with the stronger claim that restructuring occurs in the absence of literacy based on evidence of an interaction between frequency and neighbourhood density in adults with little or no knowledge of letters (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007). Psycholinguistic grain size theory also predicts that children's phonological representations will be restructured as children's vocabularies grow, but within this account phonemic representation will not appear within the lexicon until children learn the mappings between letters and sounds (Ziegler & Goswami, 2005). Another variant of the emergent view is the PRIMIR framework which proposes that words are stored across a number of multidimensional planes (Werker & Curtin, 2005). While the word form plane stores clusters of similar exemplars of words, the phonemic plane gradually extracts regularities from the word level plane leading to the formation of phonemic categories. Within PRIMIR 'phonemes' will begin to emerge early on within the phonemic plane but will not

map onto adult phonemic categories until they are ‘sharpened up’ when children learn to read (Werker & Curtin, 2005). The key theoretical accounts of phonological development outlined above make different predictions about whether or not children represent words in terms of phonemes before they become literate as summarised in Table 5.1.

Table 5.1. Predictions made by key theoretical accounts of phonological development about the emergence of phonemic representation

Theoretical account	Does phonemic representation emerge in preliterate children?	
Accessibility view ^a	✓	Phonemes are present from infancy
Lexical restructuring model ^b	Unspecified	No specific prediction made but vocabulary is emphasised as the key driver of lexical restructuring
Ventura and colleagues ^c	✓	Phonemes emerge through oral language experience alone
Psycholinguistic grain size theory ^d	✗	Phonemes emerge only after children learn about phonemes explicitly – usually when learning about letter-sound mappings
PRIMIR ^e	Not fully	Phonemes begin to emerge early but only become adult-like when children learn to read

^a e.g. Rozin & Gleitman (1977), Liberman, Shankweiler & Liberman (1989), ^b Metsala & Walley (1998), ^c Ziegler & Goswami (2005), ^d Ventura, Kolinsky, Fernandes, Querido & Morais (2007), ^eWerker & Curtin (2005).

The development of children’s phonological representations has been investigated in behavioural studies using a range of speech perception (e.g. Metsala, 1997, Garlock, Walley & Metsala, 2001; Ventura, Kolinsky, Fernandes, Querido & Morais, 2007) and similarity classification tasks (Treiman & Breaux, 1982; Storkel, 2002; Carroll & Myers, 2011; Chapters 2, 3 and 4). For example, Metsala (1997) found evidence on a gating task that children’s representations of high frequency words from dense neighbourhoods are more segmental than those of low frequency words from sparse neighbourhoods providing support for the lexical restructuring model. More recently Ainsworth and colleagues (Chapters 2 and 3) found that children’s tendency to make segmental classifications (i.e. classifications based on phonemic similarity rather than global similarity) was uniquely related to vocabulary growth and not letter-sound knowledge, consistent with the idea of lexical restructuring independent of literacy. In a related study Ainsworth et al. (Chapter 4)

tracked the development of sensitivity to phonemic versus global similarity. Both sensitivity to the number of shared phonemes between words and sensitivity to global similarity were found to increase developmentally within the first two years of school. As discussed above there are a number of empirical studies which point towards phonemic sensitivity growing in proportion to the size of the lexicon (Ainsworth, Welbourne & Hesketh, submitted, Chapters 2, 3 and 4). According to the lexical restructuring model (Metsala & Walley, 1998) which predicts a 'whole to parts' representational shift, we might also expect a corresponding decrease in sensitivity to global similarity. Interestingly however, a recent study (Chapter 4) has shown global sensitivity to be maintained and to even increase over the same developmental period. Although these behavioural studies provide support for the idea of vocabulary driven restructuring, they are not able to address questions concerning the exact mechanism by which these representations change. Computational models on the other hand allow us to probe the development of the representations directly (Mareschal & Thomas, 2007) and provide a potential window into the processes involved in PR development.

Computational models of spoken language

McClelland and Elman's TRACE model of spoken word recognition (1986) simulated the representation of phonological word forms within a connectionist framework. Within TRACE word forms are stored over three levels: features, phonemes and whole words. While connections within the same level are inhibitory, connections between levels are excitatory. This allows the simulation of both lexical competition (e.g. similar sounding words become activated as the stimulus unfolds over time) and top down influences on spoken word recognition (e.g. words are recognised more quickly if heard within a related context). The representations within TRACE consist of mock speech phonemic inputs and abstract feature vectors which assign values to different articulatory dimensions such as voiced, diffuse, etc. The TRACE model was successful in reproducing several key empirical phenomena within the speech perception literature. For example TRACE simulated categorical perception of phonemes and the interaction between feature level and word level influences on phoneme identification. TRACE was limited however by the artificial nature of the speech inputs. The majority of the findings were obtained using mock inputs rather real speech data which make a number of simplifying assumptions such as all features having equal salience and all phonemes spanning the same length of time (McClelland & Elman, 1986). TRACE is also limited by the lack of a centrally stored representation (McClelland & Elman, 1986). While TRACE involves duplication of units at each of three levels multiple times, the authors point out the need for

a 'central representational structure' (McClelland & Elman, 1986, pp.77) to allow simulation of priming effects and processing of indexical characteristics of speech such as speaker rate and accent. A further limitation of TRACE is that it does not speak to the question of how phonological representations develop over time.

Plaut and Kello's later model (1999) addresses this final limitation through the generation of a developmental model which was trained to babble, comprehend, imitate and name words. Within this model the phonological layer is a learned hidden representation which integrates information from the semantic, acoustic and articulation layers. Unlike the TRACE model, phonemic representation is not explicitly built into the model, rather it is allowed to naturally emerge within the hidden phonology layer. The model was tested at eight developmental stages on naming, repetition and comprehension. Analysis of the types of errors made by the model provided evidence of emerging segmental sensitivity within the models' phonological representations driven by oral language experience. Where the model incorrectly produced or misunderstood a word the foils were found largely to be phonologically similar to the target (Plaut & Kello, 1999). The generalizability of these results is however limited by a potential segmental bias imposed on the model by the structure of the acoustic and articulatory representations. In both cases the representations are constructed by dividing words into chunks in time and assigning values to each chunk over multiple dimensions. While the articulatory representations were constructed based on six dimensions: oral constriction, nasal constriction, place of constriction, place of constriction, tongue height, tongue backness and voicing, the acoustic representations were coded in terms of ten dimensions: formant frequencies, formant transitions, frication, burst, loudness and jaw openness. Although no segmental structure was imposed on the hidden phonology layer itself, Kello and Plaut (2004) point out that the segmental sensitivity observed within their earlier model may be due at least in part to the fact that the articulatory and acoustic input/output representations have a segmental structure.

Harm and Seidenberg (1999) simulated the development of phonological representations within a preliterate model and investigated how the representations changed when the network learnt to read. The pre-reading model was shown to encode segmental information about the sounds in words, e.g. storing 'eat' in a similar way to 'treat' and 'meat'. As the model learnt to read, the model's representations became increasingly segmental. The literate model performed better on phoneme restoration than the pre-literate and control models. Investigation of the connections weights within the attractor network showed that literacy training led to greater weight changes within segments than between segments (Harm & Seidenberg, 1999). Again, as for the Plaut &

Kello model (1999) the emergence of segmental representation may have been due in part to the segmental nature of the input representations. The input representations for the Harm and Seidenberg (1999) model were constructed using abstract dimensions to describe the key properties of the sounds in words (e.g. the level of sonorance of a phoneme). In this case the inputs were divided phonemically and so it is possible that the emerging sensitivity to phonemes observed within the model may have been partly due to the phonemic nature of the input representations.

Kello and Plaut (2004) took an initial step towards resolving this issue by training a connectionist network with real speech data which had not had any phonemic structure imposed upon it. Within this model the input and output representations were created from acoustic recordings and corresponding articulatory measurements of spoken sentences from the MOCHA database (Wrench & Hardcastle, 2000). Kello and Plaut (2004) showed that a feed forward model was able to learn the mappings between real articulatory input and real acoustic output data, but they did not investigate the way that the model learned them (i.e. the nature of the hidden representations).

The current study aimed to test the predictions made by key theoretical accounts of phonological development (Table 5.1) with a developmental computational model trained on the speech-sound mappings (between how words sound and how we say them) for a corpus of CVC words. The hidden unit representations were analysed at different time points within the training period in search of evidence of emerging phonemic representation. By using real speech and keeping the architecture and representations as theoretically neutral as possible (i.e. by not imposing phonemic structure on any part of the model) we were able to test whether or not phonemic representations emerge naturally in a preliterate model which has been exposed to oral language alone with no experience of literacy. We were also able to analyse how sensitive the model is to global similarity and whether this changes as the model matures.

Method

Model overview

A feed forward model was constructed within LENS (Light, Efficient Network Simulator, (Rohde, 2000)) with the architecture illustrated within Figure 5.1. Acoustic input representations based on real speech were fed into the model, with articulatory representations as the output. The input and output representations were connected via a hidden layer consisting of 100 units. All input units were connected to all hidden units which in turn were connected to all output units. The architecture was designed to mimic the learning of speech-sound mappings in children (i.e. the mappings between what words sound like and how we say them). The direction of the learning process (i.e. acoustics as input and speech as output) was chosen to reflect the process of word repetition which is arguably a key aspect of early language acquisition.

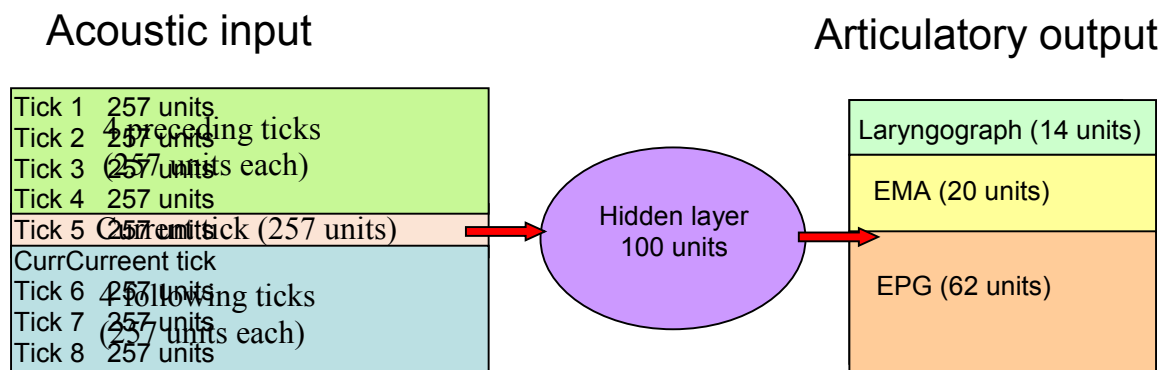


Figure 5.1. Architecture of the feed forward model.

Representations

The input representations were generated from acoustic recordings of a female speaker taken from the MOCHA database (Wrench & Hardcastle, 2000). The database consists of 460 sentences designed to provide a comprehensive range of speech sounds and coarticulation effects. Training examples were created by extracting all the CVC words from the acoustic sentence files. This was done by segmenting the power spectrum of each sentence into words using the timings for where each phoneme starts and ends provided within MOCHA. The power spectra were generated using the 'spgrambw' function within Matlab. A bandwidth of 61Hz was chosen so that the Hamming window length would equal the shortest word length within the database ('a', length 30ms). The power spectra were calculated using a Bark rather than a linear scale to reflect the nonlinear relationship between frequency and sensitivity within the human ear. The acoustic recordings were

sampled at 16000Hz and low pass filtered at 8000Hz, resulting in 257 frequency bins following the power analysis.

Temporal dynamics

To take into account the fact that articulatory gestures are dependent on quite a wide time window of the acoustic stream the representations included the 257 units for the current tick (time slice) as well as the 257 units for the four preceding and four following ticks (2313 units in total). Because articulatory gestures might extend over several time slices, a given chunk of acoustic input needs to be processed in context rather than in isolation. This arrangement can be understood in terms of the model listening to and remembering a portion of the acoustic input before the corresponding articulation is made. This ‘windowing’ through time of the continuous speech stream into contextualised chunks of input (as shown in Figure 5.1) reflects the fact that in repetition the current portion of the signal being listened is always slightly ahead of the portion of the signal that is being spoken. A similar approach was adopted within the NETtalk model (Sejnowski & Rosenberg, 1986).

An alternative approach would have been to have used a recurrent (rather than a simple feedforward) network. The use of a recurrent network where units are connected within a temporal cycle allows the model to effectively ‘remember’ sequences of information (Williams & Zipser, 1995). Given that humans use short term memory when processing speech, recurrent networks arguably provide the most realistic choice for a model of speech processing (e.g. Elman, 1990). However, recurrent networks are very complex in nature and are infamously difficult to train and interpret (e.g. Bengio, Simard & Frasconi, 1994). Given that the key focus of the current study is the analysis of the model’s hidden unit activations over training, it was important that the network architecture was sufficiently simple to allow us to make sense of how the model’s representations evolve over development. It was therefore decided that a feedforward network would be more suitable than a recurrent network on the grounds of interpretability. If a recurrent approach had been taken the model would have had to learn to perform the short term memory task as well as learning useful representations. By using a feedforward network with a “history” built into the inputs we are effectively performing the job of short term memory job for the model and allowing it to focus purely on developing useful representations. This greatly simplifies the task, but it prevents us from detecting any influence of the short term memory task on the representations themselves.

The articulatory output representations were created from three types of physiological measurement: laryngograph (LYN), electromagnetic articulograph (EMA) and electropalatograph (EPG) data. The first 14 units were generated from the LYN data which measure voicing at the larynx. The same Hamming window length was applied as for the acoustic representations, but a linear rather than a Bark scale was used. Laryngograph measurements were sampled at 16000Hz and low pass filtered at 400Hz resulting in 14 frequency bins. The next 20 units within the articulatory representations were generated from the EMA measurements which plot the positions of eight sensors within the mid-sagittal plane of the vocal tract. The EMA data consist of nine $\{x,y\}$ pairs of dimensions relating to the following positions: lower incisors relative to upper incisors; upper lips relative to upper incisors; lower lips relative to lower incisors; soft palate relative to upper incisor; absolute position of the tongue averaged over three measurements; average tongue position relative to upper incisors; three pairs based on individual tongue measurements relative to average tongue position. The final 62 units correspond to the EPG data which measure contact between the tongue and the palate across 62 sensors: where the tongue makes contact with a sensor a value of 1 is given; where no contact is made the value is zero.

The EMA and EPG data were originally sampled at 500Hz and 200Hz respectively. Both sets of data were resampled to match the sampling rate of the acoustic and LYN spectra. In order to remove extreme outliers within the acoustic LYN and EMA data, the smallest and largest 100 values within each type were set equal to the smallest and largest values respectively (a procedure also used by Kello and Plaut, 2004). The data were then normalised to a range of 0 to 1 using the minimum and maximum values calculated over the whole corpus – normalisation was unnecessary for the binary EPG data.

Training

The training examples consisted of all 534 CVC words extracted from the MOCHA database (see Appendix 12), which were presented as one batch. The weights on all connections were initially randomised to a range of $[-1\ 1]$ and then updated after each presentation of the example set using backpropagation of error signals from the output units (Rumelhart, Hinton & Williams, 1986). For each tick (time slice) the input activation values were made equal to the values specified within the acoustic representation. These were then transferred through the model via the weighted connections to create new activation values within the hidden layer and then finally through another set of weighted connections to the articulatory output units. For each output unit the sum squared error between the target output values (specified within the training examples) and the network's

output activation values was calculated and then propagated back through the model generating a set of weight derivatives, which were then used to update the connection weights. The weight updating algorithm used was ‘Doug’s momentum descent’ which is the default setting within LENS (Rohde, 2000). Within Doug’s momentum descent the weight change vector is bounded leading to greater stability during initial learning (in comparison to standard momentum descent – see Rohde, 2000 for further details). A learning rate of 0.1 was used with the weight decay set to 0.000001. Frequency effects were simulated using the ‘pseudoExampleFreq’ function, with word frequencies taken from the SUBTLEX database (Brysbaert & New, 2009). The ‘pseudoExampleFreq’ function scales the error and output unit error derivatives on each example by its frequency.

Results

How accurate is the model?

The final model (after 2000 updates) was initially tested using error thresholds of 0.1 for the multi-valued output groups (LYN and EMA) and 0.5 for the binary EPG group. The model was judged to have ‘passed’ a given example if the difference between the output values and target values were less than the thresholds stated above for all units. According to these criteria the model did not succeed on any of the examples even after 2000 updates despite showing high levels of accuracy for the LYN and EMA units (mean error per unit per tick <0.01). This was due to the relatively high error rate for the EPG units, which had a mean error per unit per tick of 0.4. In other words, while the model was very good at producing the correct LYN and EMA output patterns it was unable to get all the EPG units correct within any one example. This low pass rate for the EPG units was explained following inspection of the training examples for repeated words within the corpus (e.g. the 12 instances of ‘this’ within the database). It was found that there was a lot of variability in terms of the EPG measurements for different instances of the same word (i.e. when a single speaker says the same word within different contexts their tongue is not always making the same points of contact with the palate). A more appropriate test for the model would be to see if its outputs are as close to the targets as the targets for repeated words are to one another. Put simply, if the model deviates from the targets by no more than the variation that we see for repeated words within the training examples then we can say that the model has successfully learnt the words. The mean percentage agreement between repeated words within the training examples and the percentage agreement between outputs and target values at different developmental points within the model are shown within Table 5.2. The percentages are calculated based on the number of units which were within 0.1 of the target for the multi-valued units (i.e. LYN and EPG) and within 0.5 for the binary EPG units. As shown in Table 5.2, after 2000 updates the accuracy of the model was equal to the degree of agreement between repeated utterances within the training examples. The model was therefore judged to have successfully learnt the words by this point.

Table 5.2. Comparison of mean agreement between repeated words within the training set and mean percentage accuracy for all words at different developmental points (standard deviation given in brackets)

	LYN units	EMA units	EPG units
Mean agreement between repeated words	95.61 (4.03)	100.00 (0.07)	86.11 (10.26)
Mean accuracy at t=0	14.32 (.36)	16.62 (1.55)	48.03 (11.08)
Mean accuracy at t=100	96.26 (3.17)	91.9 (4.91)	84.12 (11.84)
Mean accuracy at t=500	96.47 (3.08)	99.39 (1.65)	85.21 (11.79)
Mean accuracy at t=1000	97.15 (2.99)	99.64 (1.16)	85.56 (11.51)
Mean accuracy at t=2000	97.31 (2.91)	99.81 (0.74)	85.91 (11.18)

How well does the model generalise its knowledge to unseen words?

In order to test how well the network is able to generalise to unseen words, another model was run which used the same parameters as the main model described above but with a training set consisting of 440 rather than 534 words. The remaining 94 words were used to test the model at different developmental time points. The model was able to apply its prior learning to unseen items as evidenced by the mean percentage accuracies displayed within Table 5.3. After 2000 updates the model was found to do equally well (within 1% percentage accuracy) on seen and unseen words.

Table 5.3. Mean percentage accuracy for 94 untrained test words at different developmental points (standard deviation given in brackets)

	Laryngograph units	EMA units	EPG units
Mean accuracy at $t=0$	20.18 (1.44)	14.86 (0.98)	55.11 (4.91)
Mean accuracy at $t=100$	94.26 (4.02)	90.46 (6.17)	83.07 (23.46)
Mean accuracy at $t=500$	94.88 (4.08)	98.96(2.26)	84.25 (13.24)
Mean accuracy at $t=1000$	94.97(4.06)	99.34 (1.61)	84.66 (12.94)
Mean accuracy at $t=2000$	95.04 (4.03)	99.70 (0.86)	85.04(12.53)

Is there evidence of emerging segmentation within the model’s hidden representations?

Once we had established that the model was capable of learning the mappings between real acoustic input and articulatory output patterns, we then set out to investigate the way that the mappings were stored. In particular we wanted to look for evidence of emerging phonemic representation within the hidden units. To do this we compared the hidden unit representations for different pairs of words and investigated whether they were related in terms of phonemic or global similarity.

The comparison of hidden unit representations was complicated by the fact that the representations are not all the same size (some words are longer than others) making the use of standard metrics such as Euclidean distance problematic. The issue was complicated further by the fact that the segmentation of individual words from the sentences within MOCHA was based on approximations (provided within the label files from the MOCHA database) of where one word starts and another begins. Because the words are extracted from continuous speech there is inevitably some overlap between the representations of consecutive words and the beginnings of words are not perfectly aligned in time (i.e. the words do not all start exactly at the first tick). To illustrate this problem a principal components analysis was conducted on the hidden unit activation values (with varimax rotation) and the factor scores for the first five principal components were plotted against

time for three different instances of the words ‘big’ and ‘can’ (Figure 5.2). Factor scores were plotted rather than individual hidden unit activation values to allow us to represent as much of the data as possible within the figure.

We can see from Figure 5.2 that the plots of different tokens of the same word (e.g. big1 and big2) look more similar in terms of overall form than tokens of different words (e.g. big1 and can1) as we would expect. There is however a lot of variation between the plots of same word tokens in terms of temporal alignment (e.g. the peaks within the plots for the three tokens of big do not all start at the same tick number), duration (the words are not all the same length in time) and the amplitude of the plots. Because we find so much variability (in terms of start point, amplitude and length) even for utterances of the same word it would be inappropriate to compare the hidden unit representations of words using a standard similarity metric such as Euclidean distance. Instead, we constructed a similarity metric designed to capture the overall form of the representation (the overall shape of the waveforms within Figure 5.2) independent of amplitude and alignment in time.

Similarity metric

A representation was devised for each set of hidden units (i.e. for each word) based on how the hidden unit activations change over time. Because of the high variability between tokens of the same word (Figure 5.2), the representation was based on the number and width of the local maxima (peaks) and minima (troughs) within the activation values rather than on their absolute size or position.

The first and second dimensions within the representation consisted of the number of peaks and troughs respectively. The number of peaks was estimated computationally using the ‘findpeaks’ function within Matlab (MathWorks, 2014). A minimum peak prominence threshold of 0.025 times the maximum amplitude was set to ensure that only major peaks were counted and very small perturbances along the waveform were ignored. Similarly troughs were found by performing ‘findpeaks’ on the inverse of the factor score plot. The third and fourth dimensions within the representations consisted of the width of each peak and trough respectively (also calculated using Matlab’s ‘findpeaks’ function (MathWorks, 20014)). Because the hidden activation plots have varying numbers of peaks and troughs, the size of the peak width and the trough width dimensions also varied. For example, if we look at example activation plots for the fifteenth hidden unit (Figure 5.3) we can see that while ‘big’ has two peaks and three troughs, ‘can’ has one peak and no troughs. The corresponding representations used to calculate the distance between these two words are illustrated within Figure 5.4.

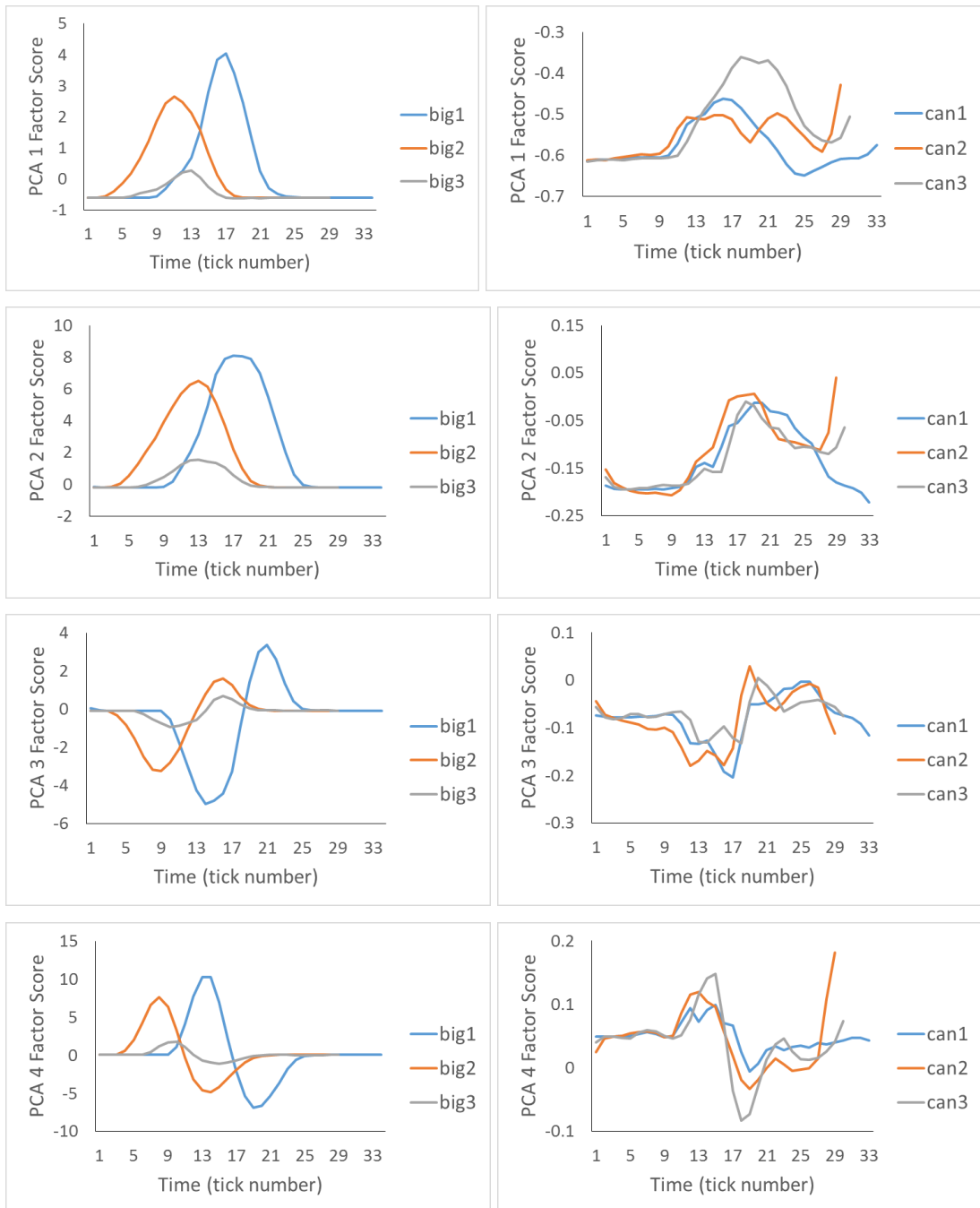


Figure 5.2. Plots of the first four principal components against time for three different tokens of the words ‘big’ (left) and ‘can’ (right). The principal components were extracted from the model’s hidden unit activation values after 2000 updates.

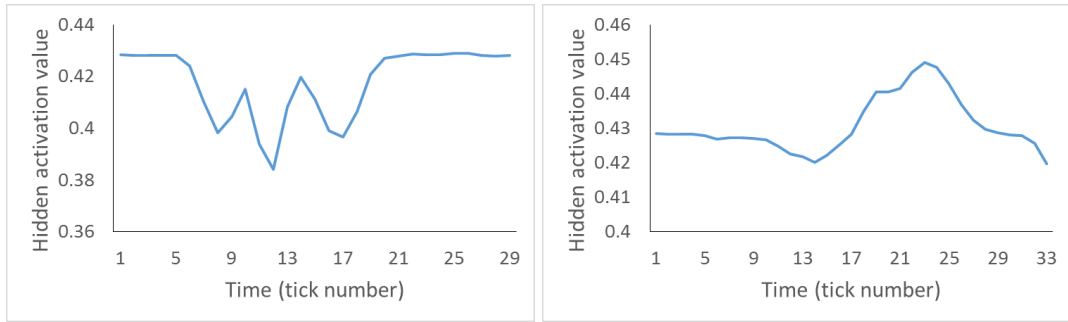


Figure 5.3. Activation plots for the hidden unit number 15. The examples shown are taken from the words ‘big’ (token 3, left) and ‘can’ (token 1, right).

For each dimension the sum square distance was calculated (for all word pairs) summing over all 100 hidden units. In cases where the number of peaks (or troughs) differed between words, only the dimensions for which both words have an entry were included within the calculation (as exemplified within Figure 5.4). The sum squared distances for each dimension were then normalised by dividing each one by the maximum distance within the corpus for that dimension. The overall ‘hidden plot distance’ was then calculated as the sum of the normalised distances for each dimension.

Hidden unit 15. Word _i = ‘big’ (token 3)		Hidden unit 15. Word _j = ‘can’ (token 1)	
No. peaks	2	No. peaks	1
No. troughs	3	No. troughs	0
Peak 1 width	1.19	Peak 1 width	8.50
Peak 2 width,	2.26		
Trough 1 width	1.92		
Trough 2 width	2.49		
Trough 3 width	2.87		

Figure 5.4. Example representations for the words ‘big’ (token 3) and ‘can’ (token 1). Only the dimensions for which both words have an entry are used to calculate the hidden plot distance as shown by the arrows.

The rationale behind the distance metric outlined above was to find a way to compare the pattern of hidden unit activations between words of different lengths which have variable amplitudes and are not temporally aligned. If the metric is successful in capturing the important differences between words, then we would expect repeated tokens of the same word (e.g. big1-big2) to have smaller inter-word distances than tokens of different words (e.g. big1-can1). In line with this prediction, Figure 5.5 shows the mean distance between different tokens of the same word to be significantly below the mean distance between different words, as confirmed by an independent samples t-test, $t(1249)=49.48, p<.0005$ (equal variances not assumed).

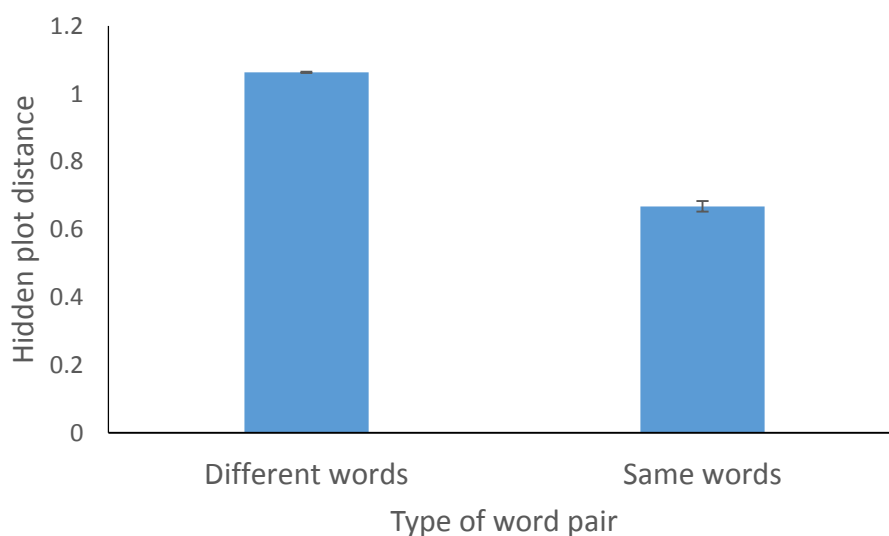


Figure 5.5. Bar chart comparing the mean hidden plot distance of different words with the mean hidden plot distance of different tokens of the same word. Error bars represent a 95% confidence interval.

Sensitivity to the number of shared phonemes and global similarity

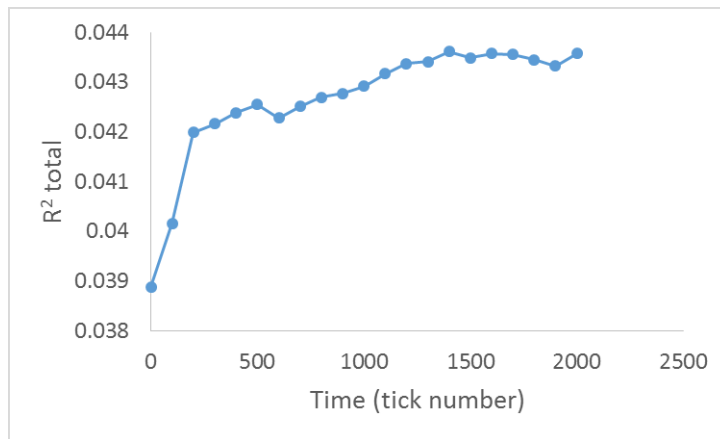
If the model is sensitive to phonemes we would expect the hidden plot distance for a given pair of words to be related to the number of shared phonemes between them, e.g. we would expect ‘can’ and ‘cat’ to have a smaller hidden plot distance than ‘can’ and ‘cup’ which in turn would have a smaller distance than ‘can’ and ‘hill’. Similarly if the model is sensitive to global similarity we would expect the hidden plot distance to be related to how close two words are globally. The global distance between words was calculated using judgement ratings collected by Singh and colleagues (Singh & Woods, 1971; Singh, Woods & Becker, 1972). This involved the concatenation of distances for corresponding phoneme pairs within the two words (as conducted by Treiman & Breaux, 1981; also Carroll & Snowling, 2001). For example the global distance between ‘head’ and ‘pet’ is calculated as the distance between /h/ and /p/, plus the distance between /ɪ/ and /ɛ/ plus the

distance between /d/ and /t/. Where one or both of the words contained phonemes not present within the rating tables, the case was omitted from the analysis – this left 89253 valid cases (pairs of words). The hidden plot distance for the final model (t=2000 updates) was found to be significantly correlated with both the number of shared phonemes ($r = -.16$, $p < .0005$) and global distance ($r = .21$, $p < .0005$). However, because the number of shared phonemes and global distance are highly correlated with each other ($r = -.92$, $p < .0005$) further analysis (see regression below) was needed to separate the potential influences of phoneme versus global similarity on hidden plot distance.

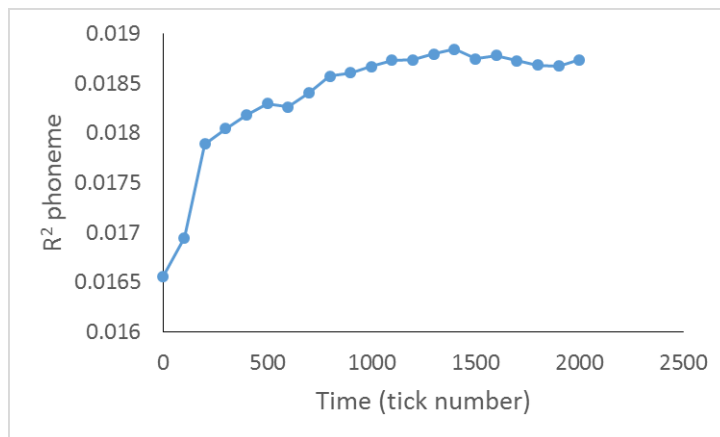
A multiple regression was performed with the number of shared phonemes and the global distance between words as the predictor variables and hidden plot distance as the outcome variable. In order to separate the amount of variance contributed by the two predictors, Johnson's method was used (Johnson, 2004) which takes into account the fact that the predictors are related. Johnson's method performs principal components analysis on the predictors to generate two orthogonal components which optimise prediction of the outcome variable. The orthogonal components are then entered into the regression model, allowing us to calculate how much of the variance is explained by each of the two predictors. Within the current study Johnson's method was applied using an automated program created by Lorenza-Seva, Ferrando & Chico (2010).

The regression analyses described above were conducted at several developmental time points to allow us to plot changes in sensitivity to both phoneme and global similarity (as measured by R^2 phoneme and R^2 global) over time. We can see from Figure 5.6 that sensitivity to both phoneme and global similarity rises over training. The hidden unit representations were found to show some segmental sensitivity prior to training as evidenced by a significant value of R^2 phoneme at t=0 (R^2 phoneme=0.017, $p < .0005$). This is due to the fact that the values within the acoustic input representations themselves are sensitive to which phonemes are present within the word (i.e. if words share more phonemes they are more similar acoustically). Pre-training sensitivity to global similarity was also found to be significant (R^2 global=0.022, $p < .0005$). Within the final model the total amount of variance explained by the two predictors (t=2000 updates) rose by 13% from its original value at t=0 moving from .035 to .044.

a)



b)



c)

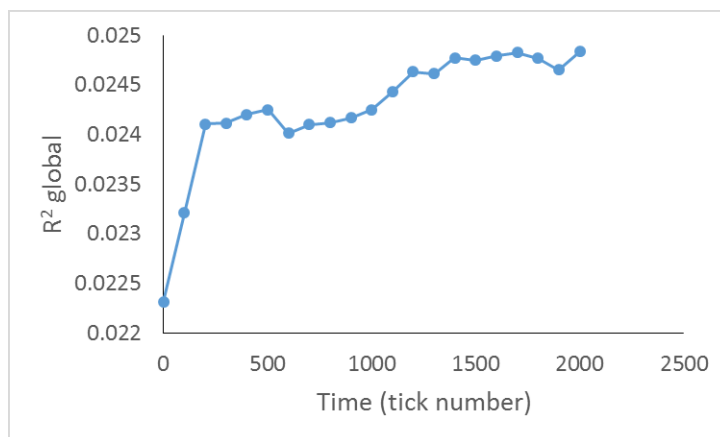


Figure 5.6. Plots against time of a) the total amount of variance in hidden plot distance explained by the regression model, b) the amount of variance explained by phoneme similarity and c) the amount of variance explained by global similarity

Discussion

The current study is the first to investigate the emergence of segmental representation in a neural network model using real speech data. A feed forward model was trained on acoustic input generated from speech recordings and articulatory output based on physiological measurements taken from the same speaker. By training the model using representations which had not had any phonemic structure imposed on them, we were able to test whether segmental representation can emerge naturally through oral language experience alone. The hidden unit representations were analysed at several developmental points by comparing their distance using a novel similarity metric (hidden plot distance) designed to capture the overall form of the representations, independent of length, temporal alignment and amplitude.

The study showed that a feed forward model is able to learn the mappings between real acoustic input and articulatory output representations. While the model found it relatively easy to learn the relationship between the acoustic inputs and both the LYN and EMA outputs, it found it much harder to learn the mappings between the acoustic data and the EPG measurements. This was due to relatively high levels of variability found within the EPG input patterns even for repeated words spoken by the same speaker.

Does phonemic representation emerge within a preliterate model?

The hidden unit representations were found to show some segmental structure even before training due to the fact that the values within the acoustic input representations themselves are sensitive to which phonemes are present within the word (i.e. if words share more phonemes they are more similar acoustically). As the model learnt the mappings between acoustic input patterns and articulatory output patterns the distance between the hidden unit representations for different words became increasingly related to the number of shared phonemes, suggesting that the model's hidden representations became more componential.

The fact that the model shows some sensitivity to phonemes from the outset lends support to the idea of early specificity proposed by the accessibility account (Rozin & Gleitman, 1977; Liberman & Shankweiler & Liberman, 1989; Bailey & Plunkett, 2002, Ballem & Plunkett, 2005, Swingley & Aslin, 2000; Swingley, 2009). However, the fact that the representations became increasingly segmental during training argues against a strong version of the accessibility view (where phonological representations are assumed to be adult-like from the outset) and is more consistent with the lexical restructuring model which proposes phonological representations to be initially underspecified, becoming increasingly componential driven by oral language experience (Metsala & Walley, 1998;

Ventura, Kolinsky, Fernandes, Querido & Morais, 2007). The simplest way to reconcile the two findings with the theoretical predictions made within Table 5.1 is to adopt PRIMIR's view (Werker & Curtin, 2005) which proposes that phonological representations contain rich phonetic detail from the outset but with phonemic categories emerging only gradually (Werker & Curtin, 2005). This view is supported by the fact that the model's sensitivity to phonemes continues to rise between 1000 to 1400 updates after the model has already learnt to say the words accurately (within 1% of the percentage agreement between repeated words at 1000 updates).

PRIMIR is also able to explain the finding that sensitivity to global similarity also increases over training (Figure 5.6). Although the term 'global similarity' is not referred to within PRIMIR, it can be thought of in terms of the word level information stored within the word form plane. Within PRIMIR 'global' or word level representations are not restructured into segmental representations. Instead information continues to be stored at all levels throughout development (Werker & Curtin, 2005). As a child gains more oral language experience the exemplar based representations within the word form plane become augmented while phoneme like categories begin to be extracted within the phoneme plane. This is in contrast with the lexical restructuring model which predicts a shift from global to segmental representation (Metsala & Walley, 1998), which we might expect to yield a decrease in global sensitivity as phonological representations become less global and more segmental. The simultaneous increase in both phoneme and global sensitivity is consistent with behavioural data in preschool children which also showed developmental rises in phoneme and global sensitivity (Chapter 4).

One of the key advantages of using a computational model to study phonological development is that we can directly control the level of literacy experience. Within the current study the model was presented with spoken words only so that any evidence of emerging segmental representation could be attributed to oral language experience. The fact that the model's representations become increasingly sensitive to the number of shared phonemes in the absence of literacy therefore goes against the idea of letter-sound knowledge (or explicit phoneme awareness gained through direct teaching) being needed for phonemes to emerge within the lexicon (Ziegler & Goswami, 2005). However, interpretation of the results is complicated by the question of how sensitive to phonemes the hidden unit representations need to be before we can argue that they are truly phonemic. In the extreme limit, we can imagine a scenario where the hidden unit activations can be entirely determined by the phonemes present in the word, i.e. words which are phonemically identical would have identical hidden unit representations. In this case all the variance in hidden plot distance would be explained by the number of shared

phonemes and we could be confident that the model was indeed storing words in terms of phonemes. Within our model we are very far from this extreme case with the number of shared phonemes explaining a significant, but small amount of the variance in hidden plot distance. There are a number of possible reasons for the large amount of variance which remains unexplained by phonemic similarity. While the first two possibilities have a theoretical basis, the final four relate to more pragmatic issues:

- 1) **Phonemic categories are emerging but are not yet adult-like** It could be that the model is beginning to extract similarities between words but that the extracted categories have yet to become fully phonemic. If this is the case it is unlikely that our model's representations would continue to become increasingly phonemic given that Figure 5.6 shows phonemic sensitivity to plateau at around 1400 updates. It may be that the model would need additional input in order for representations to become fully phonemic, for example in the form of explicit phoneme awareness training or letter-sound knowledge. This idea of phonemic categories beginning to emerge but then needing to be sharpened up by exposure to literacy is embodied within the PRIMIR framework (Werker & Curtin, 2005). A valuable question for further investigation would therefore be, 'What happens to the model's representations when we expose it to literacy?' The next step would be to add a letter-sound knowledge module to the model and see if mastery of letter-sound knowledge precipitates a steep rise in the model's sensitivity to phonemes (Werker & Curtin, 2005; Ziegler & Goswami, 2005). Similarly a phoneme awareness module could be added to see if orthography itself plays a key role or whether it is simply a mediator for explicit knowledge of phonemes (Ziegler & Goswami, 2005).
- 2) **Different levels of representation** Alternatively it could be that the model is in fact representing words phonemically but that it is also storing information at other levels of detail. For example, within PRIMIR it is proposed that contextual exemplar based information continues to be represented within the word form plane alongside an emerging phonemic plane (Werker & Curtin, 2005). If this were the case we would expect some of the variance in hidden plot distance to be accounted for by the number of shared phonemes while other variance would be accounted for by indexical and contextual information. In order to test this possibility further, work is needed to investigate ways of analysing hidden unit representations in terms of these kinds of information.

3) **Context effects** A related point to the one above but coming from a more methodological stance is the influence of context effects on the model's representations. Because the words within the corpus were spoken within sentences rather than in isolation they are susceptible to context effects. For example, in the two sentences 'Swing your arm as high as you can' and 'Of course you can have another tunafish sandwich', the word 'can' varies in terms of the degree of stress assigned to it. This causes changes not only to the length of the word but also to the vowel quality. As a result when we compare 'can' to another word within the corpus, the hidden plot distance varies depending on which sentence 'can' was taken from. For some word pairs this effect is minimal, e.g. when comparing 'can' to 'verse' the hidden plot distances differ by only 5% between the two sentence contexts. However, when comparing 'can' to 'not' there is a 42% difference in hidden plot distance depending on the context. The model is therefore sensitive not just to the number of shared phonemes but also to the effects of context on the way that the word is spoken.

The fact that the model's inputs are derived from connected speech is an advantage in the sense that children also hear words within connected speech making it more realistic than a training set derived from words spoken in isolation, however the downside is that context effects make it harder to interpret the similarity based analyses. It would be useful in the future to train the same model on words spoken in isolation to allow us to investigate the potential influence of these effects on phonological representation.

4) **Choice of similarity metric** The 'hidden plot distance' similarity metric was devised in an attempt to abstract the overall form of the changes in hidden unit activation over time (independent of length, start point and amplitude). However, it is a relatively coarse measure which may not accurately reflect the similarity between word pairs in all cases. It is possible that when attempting to extract the overall shape of the hidden unit representation we may be losing important information which might have allowed for a stronger relationship between hidden unit distance and the number of shared phonemes. This study therefore motivates further investigation into optimising the comparison of hidden unit representations of different sizes which have high levels of variability in terms of amplitude and alignment in time.

- 5) **Limited size and scope of training set** Although the model is developmental in the sense that it allow us to investigate the model's representational structure after different amounts of training, the developmental stages at which we have analysed the model's representations do not directly map onto real developmental stages during infancy/childhood. Within this study we were limited to the words available to us within the MOCHA corpus which are not likely to provide an accurate reflection of the kinds of words that a child is likely to hear, and we have not been able to taken into account the age of acquisition of the words. We are also limited by the fact that CVC words were used. This decision was made to allow us to compare words in terms of phonemic similarity without other possible confounding influences such as word length and prosodic effects being present. Again this reduces how realistic the corpus is in terms of reflecting a child's early oral language experience. Because of these restrictions on the ecological validity of the training materials it is therefore impossible to relate the development points within the model to real developmental stages. In future work, training the model with a more comprehensive and realistic training set would allow us to simulate phonological development more closely.
- 6) **Noise arising from natural variation** There is also the possibility that the data is just very noisy due to natural variation in both the acoustic inputs and articulatory outputs and that the hidden units are faithfully representing that noise. In this case most of the variation would not be useful in terms of distinguishing between different word forms.

To summarise, the relatively small amount of variance in hidden plot distance explained by phoneme similarity may be accounted for by phonemic representation beginning to emerge but needing the addition of letter-sound knowledge to become 'sharpened up'. Alternatively (or indeed in addition) the model may represent words at multiple levels of detail with the phoneme level being just one of them. The proportion of variance explained may also have been limited by issues with the choice of similarity metric, the noise associated with context effects and/or the particular training set used. In order to draw stronger conclusions about the extent to which phonemes emerge in the absence of literacy further work is needed. The addition of literacy experience to the model would be of particular theoretical interest.

Conclusion

The current work is the first computational study to show evidence of emerging phonemic representation using real speech data. Unlike previous studies (Plaut & Kello, 1999; Harm & Seidenberg, 1999), which potentially bias the network towards phonemic representation by using phonemic inputs and outputs (Kello & Plaut, 2004), the current model provides evidence that sensitivity to phonemes emerges naturally within a preliterate model where no phonemic structure has been imposed on the network architecture. The gradual emergence of phonemic representation within a preliterate model is consistent with the idea of lexical restructuring driven by oral language experience alone (Metsala & Walley, 1998; Ventura, Kolinsky, Fernandes, Querido & Morais, 2007). The fact that a large proportion of variance in the distance between the model's phonological representations remains unexplained by both phoneme and global similarity could be explained by the idea of concurrent representation at multiple levels as proposed within PRIMIR (Werker & Curtin, 2005). Further work is needed however to investigate the potential limiting effects of the training set and similarity metric used. A further avenue of particular interest would involve investigating the effect of adding literacy experience to the model.

GENERAL DISCUSSION

The overarching aim of the thesis was to investigate if and how children's phonological representations change over development. As outlined within Chapter 1, current theories of phonological development make differing predictions with regard to two key themes: accessibility versus emergence and the precipitation of the phoneme. Within chapters 2 to 5 we set out to test these predictions using both behavioural and simulation data. The following section summarises the key findings from these chapters. We then discuss the theoretical significance of these results as well as the methodological implications. In the final section we make summary conclusions along with suggestions for future work.

Summary of thesis findings

Within Chapter 2 we used novel implicit measures of phonological representation to investigate whether children's representations become increasingly segmental over time using a cross-sectional design (aged 3;2 to 5;7). As discussed in previous chapters, current models of phonological development make different predictions about whether or not children's phonological representations are adult-like from infancy and, if not, whether the emergence of phonemic representation is driven by vocabulary growth or literacy experience. Children's implicit sensitivity to shared segments between words was found to increase over development and was found to be predicted by vocabulary size but not letter-sound knowledge. Conversely, performance on explicit PR tasks was found to be predicted by letter-sound knowledge but not vocabulary size. This finding suggests that while explicit phonological awareness may be dependent on knowledge of letter-sound correspondences, segmental phonology at the representational level emerges independently.

Within Chapter 3 we followed up the youngest children (from Chapter 2) with the same measures at three further time points. The longitudinal design provided sufficient power for us to test the predictions made by the different theoretical accounts in more depth. Specifically we were able to perform regression analyses on children's performance separately for rime versus phoneme level items. This allowed us to test the different theoretical pathways illustrated within Figure 3.1. Children's implicit sensitivity to both shared rime and shared phoneme segments within words was found to be predicted by vocabulary size but not letter-sound knowledge. This is in agreement with the results of Chapter 2 but also adds stronger evidence that phoneme level representation emerges independent of orthographic knowledge.

Within Chapter 4 we aimed to separate children's sensitivity to phonemic versus global similarity. As pointed out within Chapters 1 and 4, global similarity has received

little investigation in its own right despite being implicated within both the lexical restructuring model (Metsala & Walley, 1998) and the PRIMIR framework (Werker & Curtin, 2005). Chapter 4 provides evidence from children's performance on the mispronunciation reconstruction task that both phonemic sensitivity and global sensitivity increase over development. There were however differences between the two trajectories: while global sensitivity rose early on (between young nursery and old nursery) before quickly reaching adult levels, phonemic sensitivity rose throughout the developmental period studied and was yet to reach adult levels by the end of reception.

Chapter 5 parallels the work within Chapter 4 but uses simulation rather than behavioural data. The development of phonological representations was investigated by training a neural network on the mappings between real acoustic inputs and real articulatory outputs derived from the MOCHA database (Wrench & Hardcastle, 2000). The model learnt the process of word repetition, with accuracy rates reflecting the variability found for repeated tokens of the same word. The hidden unit representations were analysed at several points during training in search of emerging sensitivity to phonemic and/or global similarity. The model was found to become increasingly sensitive to both types of similarity, in line with the behavioural results reported within Chapter 4.

Theoretical implications

Accessibility versus emergence: Are children's phonological representations adult-like from infancy?

One of the key aims of the thesis was to assess whether children's phonological representations are adult-like or whether they become gradually restructured over childhood. If children's representations are qualitatively similar to adults' then we would expect to see high performance on implicit measures of phonological representation independent of age, vocabulary size and letter-sound knowledge. Within Chapters 2 to 4, novel measures, which were designed not to require any explicit knowledge of the sounds in words, were used to probe children's implicit sensitivity to the phonological structure stored within their representations. Performance on these measures indicated that segmental representation is emerging within young nursery children but is not yet fully developed. This is evidenced by the fact that within Chapter 2 the youngest children were only just above chance on three out of five of the implicit PR measures. If the accessibility account held true we would expect children to perform above chance on all five tasks and with higher mean levels of performance (i.e. well above chance rather than just). Furthermore children's implicit PR performance was significantly predicted by vocabulary

size within both the cross-sectional (Chapter 2) and longitudinal data (Chapter 3). This suggests that segmental representation emerges gradually, driven by the pressure of a growing vocabulary (Metsala & Walley, 1998).

The findings within Chapter 5 also suggest a developmental increase in segmental sensitivity. The model's performance pre-training already showed evidence of segmental structure within its hidden unit representations. This reflects the fact that the acoustic inputs themselves have some segmental structure, i.e. words containing shared phonemes have acoustic waveforms which are more similar than those containing no shared phonemes. However, the fact that the model's phonemic sensitivity increased during training indicates that the model picked up on more than just the segmental structure within the acoustic inputs and is consistent with the idea of oral language experience causing children's phonological representations to become increasingly componential (Metsala & Walley, 1998; Ventura, Kolinsky, Fernandes, Querido & Morais, 2007). The results argue against a strong version of the accessibility account (Lieberman & Shankweiler & Lieberman, 1989; Rozin & Gleitman, 1977) for which we would expect high levels of segmental sensitivity early on, then remaining static throughout the rest of training.

The developmental rise in segmental sensitivity reported within Chapters 2, 3 and 5 is therefore consistent with both the lexical restructuring model (Metsala & Walley, 1998), which predicts a whole to parts representational shift driven by vocabulary growth, and Ventura's account (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007) which extends the LRM by making the stronger proposition that lexical restructuring occurs in the absence of literacy. It is also consistent with psycholinguistic grain size theory (Ziegler & Goswami, 2005) which predicts detail to be added to children's phonological representations at all grain sizes as they learn more words. It is important however to consider how our findings fit in with the wider literature. As discussed in Chapter 1, numerous studies have provided evidence that infants' phonological representations are stored at a similar level of detail to adults (Bailey & Plunkett, 2002, Ballem & Plunkett, 2005, Swingley & Aslin, 2000; Swingley & Aslin, 2002; Swingley, 2009, see also Ramon-Casas & Bosch, 2014 for a review). When considering how studies showing evidence of early specificity can be reconciled with the evidence of emerging segmental representation reported within this thesis, it is important to consider exactly what is being measured in both types of study. While infant mispronunciation detection studies (used to show early phonological specificity) probe the accuracy of children's phonological representations at the phone level, the implicit measures of phonological representation used within Chapters 2 to 4 probe the segmentedness of children's phonological representations at the rime and phoneme level. We can therefore account for both sets of findings by proposing that

infants' PRs are rich in detail early on but with segmental representation emerging only gradually driven by oral language experience (Werker & Curtin, 2005).

While the lexical restructuring model (Metsala & Walley, 1998) and psycholinguistic grain size theory (Ziegler & Goswami, 2005) both predict vocabulary driven restructuring (as observed within Chapters 2, 3 and 5), the extent to which they can account for infants' sensitivity to minimal contrasts remains an area of contention (e.g. Swingley, 2009; Ziegler & Goswami, 2005). On the one hand it could be argued that such infant studies contradict the LRM and psycholinguistic grain size theory, given that within both accounts phonological representations are predicted to be initially lacking in detail. As discussed in Chapter 1 however, while early specificity studies show that infants can detect novelty in the auditory stimulus, they don't necessarily tell us anything about how phonological representations are structured (Bowey & Hirakis, 2006; Ziegler & Goswami, 2005).

Although Chapters 2 and 3 set out primarily to test predictions made by the four key theoretical accounts within Table 2.1 (i.e. the accessibility account (Rozin & Gleitman, 1977; Liberman, Shankweiler & Liberman, 1989), the lexical restructuring model (Metsala & Walley, 1998), psycholinguistic grain size theory (Ziegler & Goswami, 2005) and Ventura and colleagues' account (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007), the results discussed above are consistent with two alternative accounts which frame lexical restructuring in a different way. As discussed in Chapter 1, Swingley (2009) proposes that lexical restructuring does not involve holistic representations being increasingly segmented, but rather involves probability distributions being 'sharpened up' until they become categorical. Similarly PRIMIR proposes that phonemic categories are gradually extracted from regularities within the word form plane (Werker & Curtin, 2005). Within both accounts oral language experience is predicted to drive the emergence of phonemic categories in line with our finding within Chapters 2 and 3 that segmental sensitivity is predicted by vocabulary size. The key difference between PRIMIR and Swingley's accounts on the one hand and the three emergent theories within Table 2.1 on the other, is that PRIMIR (Werker & Curtin, 2005) and Swingley's (2009) accounts both predict high levels of detail within early representations as observed in numerous infant studies (Bailey & Plunkett, 2002, Ballem & Plunkett, 2005, Swingley & Aslin, 2000; Swingley & Aslin, 2002; Swingley, 2009). In this way Swingley's account and PRIMIR allow us to reconcile aspects of both the accessibility and emergent view: within the two theories early detailed phonological representations are predicted to be present from the outset alongside emerging phonemic categories (Swingley, 2009; Werker & Curtin, 2005; also see Ramon-Casa & Bosch, 2014). This idea is supported by the fact that in our

modelling study (see Chapter 5) the model's sensitivity to phonemes continues to grow even after the accuracy threshold has been passed. The idea of detailed context-rich representations proposed within PRIMIR is also supported by Curtin's study which suggests that infants' representations contain non-criterial detail (Curtin, 2011), and by Ramon-Casas & Bosch's recent review (Ramon-Casas & Bosch, 2014) which concluded that PRIMIR (Werker & Curtin, 2005) was the theory best able to account for the findings within the early specificity literature. As discussed below, PRIMIR also has the added benefit of being able to account for our findings in relation to the role of letter-sound knowledge and developmental changes in global sensitivity. Swingley's account on the other hand makes no direct prediction about the role of orthography or global similarity (Swingley, 2009).

Precipitation of the phoneme: Does phonemic representation emerge naturally in the absence of literacy?

Within the previous section we have detailed how the current thesis supports the idea of phonemic representation emerging gradually over development. This then leads us to consider whether phonemes emerge naturally through oral language experience alone (Caravolas & Landerl, 2010; Ventura, Kolinsky, Fernandes, Querido & Morais, 2007) or whether we need to teach children about phonemes (when introducing them to the links between letters and sounds) before they begin to represent words in that way (Ziegler & Goswami, 2005). Within Chapter 2 the regression analyses conducted on the cross sectional data found implicit PR performance to be significantly predicted by vocabulary size but not letter-sound knowledge. This suggests that children's sensitivity to the shared segments within words is driven by lexical size rather than orthography. The fact that the reverse pattern of results was observed for explicit measures of phonological representation suggests that while letter-sound knowledge may not be important for segmentation of the representations themselves, it plays a key role in the development of explicit phonological awareness. The results within Chapter 2 echo those from Ventura and colleagues' adult study (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007) which concluded that lexical restructuring of phonological representations operates independent of literacy, but that significant orthographic knowledge is needed for adults to access phonemic segments during explicit phoneme awareness tasks.

Chapter 3 extends the results from Chapter 2 by analysing performance on rime versus phoneme level items separately. This more detailed analysis is motivated by the fact that current theories make different predictions about rime and phoneme level representation. While all of the theoretical accounts included within Figure 3.1 agree that

rimes will emerge as a natural linguistic unit at both a representational and a conscious level (apart from the accessibility view which does not make any direct prediction about rime), they differ in terms of whether letter-sound knowledge is needed for representation at the phoneme level to emerge (Figure 3.1). The longitudinal data presented in Chapter 3 confirm that implicit sensitivity to both rime and phoneme segments is driven by vocabulary growth and is not dependent on letter-sound knowledge. Explicit rhyme awareness (as measured by the rime level items on the explicit PR tasks) was also found to emerge in the absence of literacy, however awareness of the individual phonemes within words was found to be related to letter-sound knowledge. The fact that rhyme awareness emerges independent of orthographic knowledge is consistent with other studies within the developmental and adult literature which show both pre-reading children and illiterate adults to succeed on rhyme awareness tasks (e.g. Lenel & Cantor, 1981; MacLean, Bryant, & Bradley, 1987; de Gelder, Vroomen & Bertelson, 1993). The fact that phoneme awareness on the other hand is dependent on knowledge of letter-sound correspondences is consistent with studies which show a steep rise in phoneme awareness when children start learning to read (e.g. Cossu, Shankweiler, Liberman, Katz & Tola, 1988; Liberman, Shankweiler, Fischer & Carter, 1974; Mann & Wimmer, 2002) and studies which show adult illiterates to perform poorly on phoneme awareness tasks (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007; de Gelder, Vroomen & Bertelson, 1993). The simulation data reported within Chapter 5 provide further evidence that phonemic representation begins to emerge independent of literacy experience. The hidden plot distances between words were found to become increasingly related to the number of shared phonemes despite the fact that the model had not had any exposure to orthography or phoneme awareness training.

With reference to Figure 3.1 we can see that the pattern of results described above is consistent with both the lexical restructuring model (Metsala & Walley, 1998) and Ventura et al.'s account (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007). The results contradict psycholinguistic grain size theory however, given that it proposes that the lexicon only becomes phonemically organised after children are taught about phonemes when learning to read (Ziegler & Goswami, 2005). Again PRIMIR is able to account for our findings. Within PRIMIR, phonemic representation is proposed to emerge gradually within the phoneme plane, as regularities are identified across the word form plane (Werker & Curtin, 2005). Where PRIMIR differs from the lexical restructuring model (Metsala & Walley, 1998) and Ventura et al.'s account (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007) is that it proposes the emergent 'phonemes' to only resemble those of adults after a 'sharpening up' process when learning to read. PRIMIR therefore

predicts that the beginnings of phonemic categories will emerge before children learn about letters (as observed within Chapters 2, 3 and 5) but that representations will only become fully phonemic after children learn the links between letters and sounds. This predicted pattern of development is consistent with the fact that while the model in Chapter 5 showed increasing sensitivity to the number of shared phonemes, the percentage of variance in the hidden plot distance explained by phonemic similarity remained relatively low. One possible explanation for this is that the model's representations would not become fully phonemic unless the model gained literacy experience, allowing the emerging sensitivity to phonemes to be refined. In order for this possibility to be tested the model would need to be retrained with the addition of literacy experience in the form of an orthographic layer and/or a phoneme awareness layer. We could then investigate whether the addition of orthography and or explicit knowledge of phonemes acts to 'sharpen up' the model's sensitivity to phonemes as predicted within PRIMIR (Werker & Curtin, 2005).

There are however numerous alternative explanations for the low level of R^2 reported within Chapter 5 (see Chapter 5 for a detailed discussion). For example it could be that the (pre-literate) model is already representing words in terms of phonemes (once the phonemic sensitivity levels out at 1400 updates) but that the model is simultaneously storing information at different levels, e.g. contextual exemplar based information, also predicted within PRIMIR (Werker & Curtin, 2005). If this is the case then much of the unexplained variance might be accounted for by different levels of representation being stored within the hidden units alongside the model's emerging sensitivity to phonemes. Alternatively the large proportion of unexplained variance might be accounted for by methodological factors relating to context effects, the choice of similarity metric, the scope of the training set and noise within the input and output data.

Sensitivity to global versus phonemic similarity

A final aim of the thesis was to separate children's sensitivity to phoneme versus global similarity and to plot the trajectories of the two types of sensitivity over development. As discussed in Chapter 1, global similarity has been identified as a potential confound in measures of phonological awareness (Byrne & Fielding-Barnsley, 1993; Carroll & Snowling, 2001) but has received little attention as a separate entity of interest. In Chapter 4, children's response patterns on a mispronunciation reconstruction task were used to calculate children's sensitivity to the number of shared phonemes and the global similarity between words at different points over development. Similarly in Chapter 5 we plotted the model's sensitivity to phonemic and global similarity, this time calculated as the extent to which the distance between the hidden unit representations could be explained

by the number of shared phonemes and the global distance between words respectively. The behavioural and simulation data provided convergent results, with phonemic sensitivity and global sensitivity being found to rise over development within both studies.

Although none of the theoretical accounts discussed within this thesis make direct predictions about the developmental trajectory of global sensitivity, within the lexical restructuring model (Metsala & Walley, 1998), we would expect that as phonemic sensitivity rises, global sensitivity will fall. This is because the LRM predicts that global representations will be restructured into increasingly segmental representations. Conversely within PRIMIR we would expect that as children learn more words their sensitivity to both phonemic and global similarity will increase (Werker & Curtin, 2005). This is because PRIMIR involves multiple levels of representation which remain important and are augmented over development. While oral language experience is predicted to aid the accumulation of regularities across the word form plane leading to the extraction of phoneme-like categories within the phoneme plane, oral language experience is also predicted to add to the clustering of exemplars within the word form plane which would presumably increase children's sensitivity to similarity at the whole word (global) level. The parallel rise in phonemic/global sensitivity observed within Chapter 4 is therefore most consistent with the PRIMIR framework (Werker & Curtin, 2005).

It is interesting to note that within Chapter 4 global similarity continues to have a significant influence on the performance of adults during the mispronunciation reconstruction task. The developmental trajectory observed within Chapter 4, where global sensitivity rose early on in development (i.e. between young and old nursery children) but quickly reached adult levels, is broadly consistent with Wagenveld and colleagues' finding (Wagenveld, Segers, Alphen, van & Verhoeven, 2013) that Dutch adults were more strongly influenced by global similarity than grade 1 children. It is noted however that the grade 1 children in Wagenveld et al.'s study (aged 7;0) were much older than the oldest group of children within the current study (maximum age of 5;7) who had already reached adult levels. Wagenveld and colleagues proposed a possible explanation for the finding that adults are still heavily influenced by global similarity: while adults may store some words phonemically, they may also store some parts of words holistically (i.e. where there are sequences of phonemes which occur frequently within many words). The idea that efficient representation may not involve all word representations being fully segmental has also been suggested by Ventura and colleagues (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007) as a possible explanation for the fact that an interaction between neighbourhood density and frequency on speech processing tasks (which has been interpreted as evidence of ongoing restructuring) remains in literate adults. Alternatively

the continuing influence of global similarity on PR performance might be interpreted in terms of each word being represented both globally (i.e. exemplar based representation with the word form plane) and segmentally (as phonemic categories within the phoneme plane) within a multi-dimensional framework (Werker & Curtin, 2005). The ability of the latter possibility to readily account for the superficially contradictory findings of emergent segmental representation (Chapters 2,3 and 5) alongside evidence of early specificity within the infant literature makes PRIMIR the more appealing candidate. However further analysis at the item level to look for evidence that some words are less segmented than others in adults is needed to rule out the former possibility.

Methodological implications

Similarity judgements as a probe into the lexicon

As discussed throughout the thesis there is longstanding debate with regard to if and how children's phonological representations change over development. There are two key issues which have acted to cloud the debate: 1) ambiguity with regard to measurement at the phone versus phoneme level (Ziegler & Goswami, 2005) and 2) differences in the degree of explicit knowledge required when attempting to measure the segmentedness of children's representations. The current thesis has addressed both issues by probing children's phonological representations at the phoneme (rather than phone) level using implicit tasks which do not require any explicit knowledge of phonological structure alongside explicit measures of phonological representation.

Studies which report evidence of early specificity measure the accuracy of children's phonological representations at the phone level. While they provide compelling evidence that infants' representations are highly detailed they may not be adult-like in terms of how segmented they are. Previous studies which have failed to find a relationship between vocabulary size and infants' ability to detect mispronunciations (Swingley & Aslin, 2000; Werker, Fennell, Corcoran & Stager, 2002) have been interpreted as evidence that early representations are adult-like and that no restructuring takes place. However, within Chapters 2 and 3 where we have used measures designed to probe phoneme level segmentation rather than phone level accuracy we find clear evidence of lexical restructuring. The current thesis therefore highlights the need to differentiate between the two types of measurement when investigating children's phonological representations.

The second issue relates to the level of explicit knowledge that we require when attempting to measure the segmentedness of children's phonological representations. As shown within Table 2.1 the key theoretical accounts make different predictions about

implicit versus explicit measures of phonological representation. It is therefore important when attempting to test these predictions that we contrast performance on the two types of PR measure. While previous studies have focussed on either implicit (e.g. Storkel, 2002; Carroll & Myers, 2011) or explicit (e.g. Anthony, Lonigan, Driscoll, Phillips & Burgess, 2003; Duncan, Cole, Seymour & Magnan, 2006) measures of phonological representation (the latter often referred to as phonological awareness tasks), Chapters 2 and 3 have contrasted implicit and explicit measures of phonological representation within the same sample. Within Chapters 2 and 3 we have shown that while performance on tasks requiring an explicit awareness of phonological segments is dependent on letter-sound knowledge, performance on implicit PR tasks reveals segmental phonology which is independent of orthography. These findings underline the need to make PR tasks as implicit as possible when attempting to probe the level of segmentation of the representations themselves.

A further note on the implications of the similarity paradigm used within this thesis is that Chapter 4 represents the first time that sensitivity to phonemes versus global similarity has been separated within a developmental design. Although Wagensveld et al. (Wagensveld, Segers, Alphen, van & Verhoeven, 2013) compared the influence of global similarity on the judgements of children of different ages and also of adults, global similarity was confounded with phonemic similarity (as discussed in Chapter 4). In Chapter 4 we eliminate this confound by manipulating items carefully within the mispronunciation reconstruction task so that we can look at the influence of global similarity when the number of shared phonemes is held constant, and the influence of the number of shared phonemes when global similarity is held constant. In this way we were able to separate the two types of sensitivity and provide robust evidence of a developmental rise in the influence of both global and phonemic similarity on children's similarity judgements.

The separation of phonemic versus global sensitivity within this thesis represents a significant advance on previous studies where the two similarity relations have been confounded (Ainsworth, Welbourne & Hesketh, submitted; Wagensveld, Segers, Alphen, van & Verhoeven, 2013). However, it is important to note that the method used to operationalise global similarity within the current work has two key limitations (as discussed in detail within Chapter 4). Firstly, global similarity values are derived from adult judgement data which are inherently subjective (Kessler, 2005). Secondly, segmenting words into phonemes and then summing over the distances between corresponding phoneme pairs is arguably inappropriate when attempting to measure similarity in 'global' terms. In particular, this concatenative method ignores the effects of coarticulation which are likely to have a significant impact on inter-word similarity. This

thesis therefore motivates the development of a more refined measure of global similarity to address these limitations.

Connectionist models of spoken language using real speech data

In Chapter 5 we found evidence of emerging sensitivity to phonemes without imposing an a priori segmental structure on the network's architecture or on its input and output representations. This builds on the previous work of Kello and Plaut (2004) who demonstrated the feasibility of using real articulatory input and real acoustic output representations derived from the MOCHA database (Wrench & Hardcastle, 2000). We extended this work by training a similar network on real speech data in the opposite direction (i.e. with acoustic inputs and articulatory outputs) to simulate word repetition. We then investigated changes in the model's hidden unit representations over time. The fact that sensitivity to both phonemic and global similarity rose over training has important theoretical implications (as discussed above). This finding also has important methodological implications as it demonstrates the use of real speech data as a viable and potentially more ecologically valid method of assessing the natural emergence of phonemic representation in comparison with previous models which have been biased towards phonemic representation because of the segmental nature of either the architecture (Harm & Seidenberg, 1999) or the input/output representations (Plaut & Kello, 1999).

As discussed in Chapter 5 the fact that real speech was used presented us with a number of methodological challenges. Firstly the use of real speech led to high levels of variability within the input and output representations. Because of this, traditional measures of accuracy (e.g. where all units across ticks need to be correct within a particular threshold) would have inappropriately concluded that the model had not learnt the mappings. This issue was addressed by applying a new set of accuracy criteria which required the model to be correct to within the degree of agreement seen between repeated tokens of the same word. This may be a useful method to be adopted in future computational studies where real speech data which is inherently more variable than theoretically driven input/output representations are used.

The other key challenges involved in using real speech data relate to the calculation of hidden unit distance when looking for evidence of emerging segmental structure. As discussed in detail in Chapter 5, comparison of the hidden unit representations for different words was complicated by the fact that even repeated tokens of the same word differed in terms of duration, amplitude and temporal alignment. In an attempt to address this problem we constructed a novel similarity metric designed to capture the overall form of the hidden unit plots, independent of duration, amplitude and alignment. This metric represents an

important first step in terms of solving some of the methodological barriers involved in using real speech data. Use of this measure allowed us to detect evidence of emerging segmental structure within a network which was designed to be as theoretically neutral as possible. The metric is however relatively coarse which (as discussed above) might partly explain the low levels of R^2 reported within the regression analyses in Chapter 5. The current thesis motivates further work to refine the metric so that it captures the similarities between repeated tokens of the same word more accurately.

Implications for educational practice

This thesis also has implications for the early identification of phonological difficulties in children. Although the cause of developmental dyslexia remains a controversial topic, some authors have suggested that the quality of children's phonological representations may be implicated (Fowler, 1991; Ramus, 2003; Snowling, 2000; Stein & Walsh, 1997; Wolf, O'Rourke, Gidney, Lovett, Cirino & Morris, 2002; although see Dickie, Ota & Clark, 2013 for evidence to the contrary). The phonological deficit hypothesis proposes that the phonological representations of children with dyslexia may be underspecified in comparison with those of children without dyslexia (Ramus, 2003; Snowling, 2000). This hypothesis arose from the well-established observation that dyslexic children consistently underperform relative to their peers on phonological awareness tasks (Bruck, 1992; Fawcett & Nicolson, 1995; Pratt & Brady, 1988; Wilson & Lesaux, 2001). If the quality of phonological representations is a potential marker for children at risk of phonological difficulties and related problems when learning to read, it is therefore important that we have measures which can probe children's representations at an early age (e.g. Claessen, Heath, Fletcher, Hogben & Leitao, 2009).

Previous studies investigating the early identification of phonological impairment have tended to use mispronunciation detection tasks (e.g. Claessen, Heath, Fletcher, Hogben & Leitao, 2009) or measures of phonological awareness (e.g.; Catts, Fey, Zhang & Tomblin, 2001; Hogan, Catts & Little, 2005; Lonigan, Burgess & Anthony, 2000). The findings within the current thesis present important implications however for the use of these measures. We have demonstrated that although phone level measures used in infant studies suggest adult-like levels of detail in children's early representations, children's performance on implicit PR measures at the phoneme level suggest that phonemic representation emerges only gradually. When attempting to measure a phonological deficit which might impair a child's chance at later reading success, it is representation at the phoneme level which is of interest (Fowler, 1991). Given that mispronunciation detection tasks measure accuracy at the phone level (i.e. children's ability to detect small phonetic

changes) rather than segmentation at the phoneme level (i.e. the extent to which children store words in terms of abstract phonemic categories, Ziegler & Goswami, 2005), they may not therefore be suitable for picking up on the differences in PR segmentedness proposed to predict problems with phoneme awareness and later reading performance (Fowler, 1991).

The second type of measure often used to identify children at risk of later reading performance – phonological awareness – is limited in terms of the age at which it can be diagnostically useful. Although phonological awareness tasks have been shown to be effective indicators of later reading skills in children (Catts, Fey, Zhang & Tomblin, 2001; Hogan, Catts & Little, 2005; Lonigan, Burgess & Anthony, 2000), the findings within Chapters 2 and 3 suggest that phonological awareness performance at the phoneme level (i.e. phoneme awareness tasks) is dependent on letter-sound knowledge. When a child scores poorly on a phoneme awareness task it is therefore difficult to assess whether this is because they have an underlying phonological deficit or because they have yet to gain the letter-sound knowledge required to access the task. Poor performance in young children could also be due to the meta-cognitive demands of the task being too high. The implicit measures of PR used here provide an alternative method for assessing the precursors of phonological difficulties, with two key advantages. Firstly the tasks measure phoneme rather than phone level representation, which as discussed above is the level predicted to be related to reading difficulties (Fowler, 1991). Secondly the tasks do not require any explicit knowledge of phonological structure. This means that they are less demanding meta-cognitively and are not dependent on orthographic knowledge. The tasks therefore have the potential to identify the markers of phonological impairment earlier than has been previously possible. Given the importance of early intervention when supporting children with reading difficulties (Strickland, 2002; DCSF, 2009), further investigation into the feasibility of using these measures for early identification of phonological difficulties is recommended for future work.

The computational work within Chapter 5 also has potential implications for educational practice. As mentioned above an interesting next step would involve adding an orthography and/or phoneme awareness layer to the model to see if its representations become increasingly phonemic. While the outcome of such work would be of theoretical interest (as discussed above) in terms of providing further tests of predictions made by theoretical accounts of PR development, it could also be important in terms of the teaching of early reading. One of the current areas of debate relates to exactly when graphemes should be introduced. While government guidance on the teaching of phonics suggests that we should only teach children about letters once they have developed phoneme awareness

(e.g. by oral segmenting and blending the sounds within words (PNS, 2007)), others have argued that introducing letter knowledge early on can help to scaffold children's development of phoneme awareness (Hohn & Ehri, 1983; Adams, Treiman & Pressley, 1998) and might actually be needed before they can acquire it (Mann & Wimmer, 2002; Castles & Coltheart, 2004; Lukatela, Carello, Shankweiler & Liberman, 1995; Morais, Bertelson, Cary & Alegria, 1986; Adrian, Alegria & Morais, 1995). By training different versions of the model with letter knowledge and phoneme awareness added at different stages, we could investigate the optimal order for developing these key pre-reading skills. Given that previous modelling work has shown the order of knowledge acquisition to often be more important than the age of knowledge acquisition (Thomas & Knowland, 2009), this could be an important area for future work.

Conclusions and future directions

We have presented evidence that, while explicit knowledge of the phonemes within words is dependent on letter-sound knowledge, the emergence of phonemic sensitivity as measured on implicit measures of phonological representation occurs independently of orthographic knowledge. These findings are consistent with theoretical accounts which emphasise vocabulary growth as the key driver behind the emergence of segmental structure (Metsala & Walley, 1998) and with Ventura's study which found evidence of lexical restructuring alongside low levels of phoneme awareness within an adult illiterate sample (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007). The results are also consistent with the PRIMIR framework which proposes rich levels of detail within children's word level representations alongside an emerging phoneme plane which is driven initially by oral language experience but becomes refined when children learn to read (Werker & Curtin, 2005). Within Chapters 4 and 5 we present convergent evidence from both behavioural and simulation data that children's sensitivity to both phonemic and global similarity increases over development. This pattern of results is most consistent with PRIMIR which predicts that representations at both the word and phoneme level will continue to be augmented over development. PRIMIR is also the theoretical account which can most easily reconcile the evidence of early phonological specificity within the infant literature (Bailey & Plunkett, 2002, Ballem & Plunkett, 2005, Swingley & Aslin, 2000; Swingley & Aslin, 2002; Swingley, 2009) with evidence of emerging phonemic representation presented within Chapters 2, 3 and 5. PRIMIR also provides at least a partial explanation for the relatively low levels of variance accounted for within the regression analyses in Chapter 5 – it could be that the hidden units were representing information at

the phoneme level but were also representing contextual information as predicted by PRIMIR (Werker & Curtin, 2005).

This thesis has tested predictions made by some of the key theoretical accounts of phonological development and has identified the PRIMIR framework (Werker & Curtin, 2005) as the account which is most consistent with both our findings and the wider developmental literature. The current work has also underlined the need to distinguish between measurements of children's phonological representations at the phoneme versus the phone level (Ziegler & Goswami, 2005) when looking for evidence of emerging segmental representation. Furthermore, Chapters 2 and 3 have highlighted the need to contrast explicit PR measures with implicit PR measures which have the potential to pick up on segmental phonology at an earlier age and without the need for orthographic knowledge. This has potential implications for the early identification of phonological impairments which have been shown to predict later chances of reading success (Wagner, Torgeson & Rashotte, 1994). This thesis therefore motivates further investigation into whether the measures might be used as a tool to identify children at risk of later reading difficulties.

The computational study within Chapter 5 has shown that segmental representation emerges naturally within a preliterate model which does not have any segmental structure engineered into its inputs/outputs or its architecture. The metric used to analyse the hidden unit representations represents an important first step towards tackling the methodological issues associated with using real speech data. Further work could involve modification of this metric to maximise how well it captures the similarities in overall form between the hidden unit representations for different words. Other interesting avenues for future research include the addition of an orthographic module and/or phoneme awareness module to see if this leads to refinement of the model's representations as predicted by PRIMIR (Werker & Curtin, 2005). It would also be interesting to bring the behavioural and computational studies closer together by 'testing' the model using similar tasks to those given to the children. In this way we could provide more robust evidence of the developmental trajectories for phoneme versus global sensitivity.

References

- Adams, M. J., Treiman, R., & Pressley, M. (1998). Reading, writing, and literacy. In I. E. Sigel & K.A. Renninger (Eds.), *Handbook of child psychology, vol.4: Child psychology in practice* (pp. 275-355). New York: Wiley.
- Adrian, A.M., Alegria, J., & Morais, J. (1995). Metaphonological abilities of Spanish illiterate adults. *International Journal of Psychology, 30*, 329-353.
- Ainsworth, S. L., Welbourne, S., & Hesketh, A. (submitted). *Lexical restructuring in preliterate children: evidence from novel measures of phonological representation.*
- Altvater-Mackensen, N., Van der Feest, S., & Fikkert, P. (2013). Asymmetries in early word recognition: The case of stops and fricatives. *Language Learning and Development, 10*, 149-178.
- Anthony, J. L., Lonigan, C. J., Driscoll, K., Phillips, B. M., & Burgess, S. R. (2003). Phonological sensitivity: A quasi-parallel progression of word structure units and cognitive operations. *Reading Research Quarterly, 38*, 470-487.
- Bailey, T., & 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 & Plunkett, K. (2005). Phonological specificity in children at 1;2. *Journal of Child Language, 32*, 159-173.
- Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks, 5*, 157-166.
- Ballinger, G. A. (2004). Using Generalized Estimating Equations for Longitudinal Data Analysis. *Organizational Research Methods, 7*, 127-150.
- Bowey, J. A., & Hirakis, E. (2006). Testing the protracted lexical restructuring hypothesis: The effects of position and acoustic-phonetic clarity on sensitivity to mispronunciations in children and adults. *Journal of Experimental Child Psychology, 95*, 1-17.

- Bruck, M. (1992). Persistence of dyslexics' phonological awareness deficits. *Developmental Psychology, 28*, 874–886.
- Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*, 977-990.
- Byrne, B. and Fielding-Barnsley, R. (1993). Evaluation of a program to teach phonemic awareness to young children: A one year follow up. *Journal of Educational Psychology, 85*, 104-111.
- Caravolas, M., & Landerl, (2010). The influences of syllable structure and reading ability on the development of phoneme awareness: A longitudinal cross-linguistic study. *Scientific Studies of Reading, 14*, 464-484.
- Carroll, J. M. & Myers, J. M. (2011). Spoken word classification in children and adults. *Journal of Speech, Language and Hearing Research, 54*, 127-147.
- 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. and Snowling, M. J. (2004). Language and phonological skills in children at high risk of reading difficulties. *Journal of Child Psychology, 45*, 631–640.
- Castles, A. & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition, 91*, 77-111.
- Catts, H. W., Fey, M. E., Zhang, X. & Tomblin, J. B. (2001). Estimating the risk of future reading difficulties in kindergarten children: a research-based model and its clinical implementation. *Language Speech and Hearing Services in Schools, 32*, 38–50.
- Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America, 95*, 1570–1580.

- Claessen, M., Heath, S., Fletcher, J., Hogben, J. & Leitão, S. (2009) Quality of phonological representations: a window into the lexicon?, *International Journal of Language & Communication Disorders*, 44, 121- 124.
- Cole, R. A., & Perfetti, C. A. (1980). Listening for mispronunciations in a children's story: the use of context by children and adults. *Journal of Verbal Learning and Verbal Behavior*, 19, 297-315.
- Connolly, J. H. (1997). Quantifying target-realization differences .1. Segments. *Clinical Linguistics & Phonetics*, 11, 267-287.
- Cossu, G., Shankweiler, D., Liberman, I. Y., Katz, L., & Tola, G. (1988). Awareness of phonological segments and reading ability in Italian children. *Applied Psycholinguistics*, 9, 1–16.
- Curtin, S. (2011). Do newly formed word representations encode non-criterial information? *Journal of Child Language*, 38, 904-917.
- DCSF (2009). The National Strategies Primary. *Every Child a Reader – the layered approach*. London: DCSF.
- de Gelder, B., Vroomen, J., & Bertelsen, P. (1993). The effects of alphabetic reading competence on language representation in bilingual Chinese subjects. *Psychological Research – Psychologische Forschung*, 55, 315-321.
- Department of Education. (2011). *Supporting Families in the Foundation Years*. London: Department of Education.
- Dickie, C., Ota, M., & Clark, A. (2013). Revisiting the phonological deficit in dyslexia: Are implicit non-orthographic representations impaired? *Applied Psycholinguistics*, 34, 649-672.
- Duncan, L. G., Cole, P., Seymour, P. H. K., & Magnan, A. (2006). Differing sequences of metaphonological development in French and English. *Journal of Child Language*, 33, 369-399.
- Duncan, T. E., & Duncan, S. C. (2004). Introduction to Latent Growth Curve Modeling, *Behavior Therapy*, 35, 333-363.

- Dunn, L. M., Dunn, D. M., Sewell, J., Styles, B., Bryzaska, B., Shamsan, Y., & Burge, B. (2009). The British Picture Vocabulary Scale. London: GL Assessment Ltd.
- Elbro, C., & Petersen, D. K. (2004). Long-term effects of phoneme awareness and letter Sound training: An intervention study with children at risk for dyslexia. *Journal of Educational Psychology, 96*, 660–670.
- Elman, J. L. (1990). Finding Structure in Time, *Cognitive Science, 14*, 179-211.
- Fawcett, A. J., & Nicolson, R. (1995). Persistence of phonological awareness deficits in older children with dyslexia. *Reading and Writing, 7*, 361–376.
- Ferguson, C. A. & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language, 51*, 419-439.
- Fikkert, P. (2010). Developing representations and the emergence of phonology: evidence from perception and production. *Laboratory Phonology, 10*. 227-261.
- Fowler, A. E. (1991). How early phonological development might set the stage for phonological awareness. In S. Brady & D. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp.97-117). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Fowler, C. A., & Galantucci, B. (2005). The relation of speech perception and speech production. In D. B. Pisoni & R. E. Remez (Eds.), *The handbook of speech perception* (pp. 633-652). New York: Basil Blackwell.
- Foy, J. G., & Mann, V. A. (2009). Effects of onset density in preschool children: Implications for development of phonological awareness and phonological representation. *Applied Psycholinguistics, 30*, 339–361.
- Frith, U. (1998). Literally changing the brain. *Brain, 121*, 1011-1012.
- Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighbourhood density effects on spoken word recognition by children and adults. *Journal of Memory and Language, 45*, 468-492.
- Grainger, J., & Ferrand, L. (1994). Phonology and orthography in visual word recognition: Effects of masked homophone primes. *Journal of Memory and Language, 33*, 218-

- Harm, M. W., McCandliss, B. D., & Seidenberg, M. S. (2003). Modeling the successes and failures of interventions for disabled readers. *Scientific Studies of Reading, 7*, 155-182.
- Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review, 106*, 491-528.
- Hazen, V., & Barrett, S. (2000). The development of phonemic categorization in children aged 6-12. *Journal of Phonetics, 24*, 377-396.
- Hogan, T., Catts, H. and Little, T. (2005). The relationship between phonological awareness and reading: implications for the assessment of phonological awareness. *Language, Speech, and Hearing Services in Schools, 36*, 285-293.
- Hohn, W. E., & Ehri, L. C. (1983). Do alphabet letters help prereaders acquire phonemic segmentation skill. *Journal of Educational Psychology, 75*, 752-762.
- Hulme, C. & Snowling, M. (2009). *Developmental cognitive disorders*. Oxford, England: Blackwell/Wiley.
- Johnson, J. W. (2000). A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate Behavioral Research, 35*, 1-19.
- Juszyk, P. W., Goodman, M. B., & Baumann, A. (1999). Nine-month-olds' attention to sound similarities in syllables. *Journal of Memory and Language, 40*, 62-82.
- Kello, C. T. & Plaut, D. C. (2004). A neural network model of the articulatory-acoustic forward mapping trained on recordings of articulatory parameters. *Journal of the Acoustical Society of America, 116*, 2354-2364.
- Kessler, B. (2005). Phonetic comparison algorithms. *Transactions of the Philological Society, 103*, 243-260.
- Lenel, J. C., & Cantor, J. H. (1981). Rhyme recognition and phonemic perception in young children. *Journal of Psycholinguistic Research, 10*, 57-68.

- Li, P., & MacWhinney, B. (2002). PatPho: a phonological pattern generator for neural networks. *Behavior Research Methods, Instruments, and Computers*, 34, 408–415.
- Lieberman, A. M. & Mattingly, I. G. (1989). A specialization for speech perception. *Science*, 243, 489-494.
- Lieberman, I. Y., Shankweiler, D., Fischer, F. W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, 18, 201–212.
- Lieberman, I. Y., Shankweiler, D. & Liberman, A. M. (1989). The alphabetic principle and learning to read. In D. Shankweiler & I. Y. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 1-33). Ann Arbor: University of Michigan Press.
- Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: evidence from a latent-variable longitudinal study. *Developmental Psychology*, 36, 596-613.
- Lorenzo-Seva, U.; Ferrando, P.J.; Chico, E. (2010). Two SPSS programs for interpreting multiple regression results. *Behavior Research Methods Instruments & Computers*, 42, 29-35.
- Lotto, A. J. & Holt, L. L. (2000). The illusion of the phoneme. In S. J. Billings, J. P. Boyle, & A. M. Griffith (Eds.), *Chicago Linguistic Society, Volume 35: The Panels*. Chicago: Chicago Linguistic Society.(pp. 191-204).
- Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19, 1-36.
- Lukatela, K., Carello, C., Shankweiler, D., & Liberman, I. Y. (1995). Phonological awareness in illiterates: Observations from Serbo-Croatian. *Applied Psycholinguistics*, 16, 463-487.
- MacLean, M., Bryant, P. E., & Bradley, L. (1987). Rhymes, nursery rhymes and reading in early childhood. *Merrill-Palmer Quarterly*, 33, 255-282.

- MacWhinney, B. (2010). Computational models of child language learning: an introduction. *Journal of Child Language*, *37*, 477-485.
- Mann, V., & Wimmer, H. (2002). Phoneme awareness and pathways into literacy: A comparison of German and American children. *Reading and Writing*, *15*, 653-682.
- Mareschal, D. & Thomas M. S. C. (2007) Computational modeling in developmental psychology. *IEEE Transactions on Evolutionary Computation*, *11*, 137-150.
- MathWorks (2014). *MATLAB and Statistics Toolbox Release 2014b*, The MathWorks, Inc., Natick, Massachusetts, United States.
- McCandliss, B. D., Beck, I., Sandak, R., & Perfetti, C. (2003). Focusing attention on decoding for children with poor reading skills: Design and preliminary tests of the word building intervention. *Scientific Studies of Reading*, *7*, 75–104.
- McCandliss, B. D., Beck, I., Sandak, R., & Perfetti, C. (2003). Focusing attention on decoding for children with poor reading skills: Design and preliminary tests of the word building intervention. *Scientific Studies of Reading*, *7*, 75–104.
- Merriman, W. E. & Schuster, J. M. (1991). Young children's disambiguation of object name reference. *Child Development*, *62*, 1288-1301.
- Metsala, J. L. (1997). An examination of word frequency and neighbourhood density in the development of spoken –word recognition. *Memory and Cognition*, *25*, 47-56.
- Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In L. C. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Hillsdale, NJ: Erlbaum.
- Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology*, *91*, 3-19.
- Miller, J. L., & Eimas, P. D. (1996). Internal structure of voicing categories in early infancy. *Perception and Psychophysics*, *58*, 1157–1167.
- Morais, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy training and speech

- segmentation. *Cognition*, 24, 45-64.
- Muneaux, M., & Ziegler, J. C. (2004). Locus of orthographic effects in spoken word recognition: Novel insights from the neighbour generation task. *Language and Cognitive Processes*, 19, 641-660.
- Nittrouer, S. (2002). Learning to perceive speech: How fricative perception changes, and how it stays the same. *Journal of the Acoustical Society of America*, 112, 711–719.
- Ofsted (2014, May 12). Data Dashboard. Retrieved from <http://dashboard.ofsted.gov.uk/>.
- Pattamadilok, C., Knierim, I., Kawabata Duncan, K. J., and Devlin, J. T. (2010). How does learning to read affect speech perception? *Journal of Neuroscience* 30(25): 8435-8444.
- Plaut, D. C., McClelland, J. L., Seidenberg, M., & Patterson, K. E. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103, 56–115.
- Plaut, D. C., and Kello, C. T. (1999). The emergence of phonology from the interplay of speech comprehension and production: A distributed connectionist approach, in B. MacWhinney (Ed.), *The Emergence of Language* (pp381-415). edited Erlbaum, Mahweh, NJ.
- PNS (2007). *Letters and Sounds: Principles and Practice of high quality phonics*. London: DfES.
- Pratt, A., & Brady, S. A. (1988). Relation of phonological awareness to reading disability in children and adults. *Journal of Educational Psychology*, 80, 319–323.
- Ramon-Casas, M & Bosch, L. (2014). Consonants, vowels and levels of specification in the phonological representations of the first lexicon: a review. *Anales de Psicología*, 30, 703-715.
- Ramus, F. (2003). Dyslexia: Specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology*, 13, 212–218.
- Read, C., Zhang, Y., Nie, H., & Ding, B. (1986). The ability to manipulate speech sounds depends on knowing alphabetic spelling. *Cognition*, 24, 31-44.

- Renfrew, C. (1997). *Word Finding Vocabulary Test*. Speechmark.
- Rohde, D.L.T. (2000). *Lens, The light, efficient network simulator*.
<http://tedlab.mit.edu/~dr/Lens/>.
- Rozin, P. & Gleitman, L. R. (1977). The structure and acquisition of reading II: the reading process and the acquisition of the alphabetic principle. In A. Reber & D. Scarborough (Eds.), *Toward a psychology of reading* (pp. 55-141). Hillsdale, NJ: Erlbaum.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors, *Nature*, 323, 533–536.
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word Recognition and naming. *Psychological Review*, 96, 523–568.
- Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology-Human Learning and Memory*, 5, 546-554.
- Sejnowski, T. J. & Rosenberg, C. R. (1986). *NETtalk: A parallel network that learns to read aloud*, Johns Hopkins University Department of Electrical Engineering and Computer Science Technical Report 86/01.
- Shepp, B. E., & Swartz, K. B. (1976). Selective attention and the processing of integral and nonintegral dimensions: A developmental study. *Journal of Experimental Child Psychology*, 22, 73-85.
- Singh, S., & Woods, D. R. (1971). Perceptual structure of 12 American english vowels. *Journal of the Acoustical Society of America*, 49, 1861-1866.
- Singh, S., Woods, D. R., & Becker, G. M. (1972). Perceptual structure of 22 prevocalic english consonants. *Journal of the Acoustical Society of America*, 52, 1698-1713.
- Smith, L. & Kemler, D. G. (1977). Developmental trends in free classification: evidence for a new conceptualization of perceptual development. *Journal of Experimental Child Psychology*, 24, 279-298.
- Snowling, M. J. (2000). *Dyslexia (2nd ed.)*. Oxford: Blackwell.

- 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.
- Stein, J., & Walsh, V. (1997). To see but not to read: The magnocellular theory of dyslexia. *Trends in Neurosciences*, *20*, 147–152.
- Stevens, K. N. (1998). *Acoustic phonetics*. Cambridge, MA/London: MIT Press.
- Storkel, H. L. (2002). Restructuring similarity neighborhoods in the developing mental lexicon. *Journal of Child Language*, *29*, 251-274.
- Storkel, H.L. & Hoover, J.R. (2010). An online calculator to compute phonotactic probability and neighborhood density on the basis of child corpora of spoken American English. *Behavior research methods*, *42*, 497-506.
- Strickland, D.S. (2002). The Importance of Effective Early Intervention. In A.E. Farstrup, & S. Samuels (Eds.), *What Research Has to Say About Reading Instruction* (pp. 69-86). Newark, DE: International Reading Association.
- Swingle, D. (2007). Lexical exposure and word-form encoding in 1.5-year-olds. *Developmental Psychology*, *43*, 454-464.
- Swingle, D. (2009). Onsets and codas in 1.5-year-olds' word recognition. *Journal of Memory and Language*, *60*, 252-269.
- Swingle, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*, 147-166.
- Swingle, D. and Aslin, R.N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, *13*, 480-484.
- Swingle, D., & Aslin, R.N. (2007). Lexical competition in young children's word learning. *Cognitive Psychology*, *54*, 99-132.
- Thomas, M. S. C. & Knowland, V. (2009). Sensitive periods in brain development: Implications for education policy. *European Psychiatric Review*, *2*, 17-20.
- Treiman, R. & Baron, J. (1981). Segmental analysis ability: Development and relation to reading ability. In G.E. MacKinnon & T. G. Waller (Eds.), *Reading research: Advances in theory and practice* (Vol 3, pp.159-198). New York: Academic Press.

- 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*, 569-598.
- Treiman, R. (1992). The role of intrasyllabic units in learning to read and spell. In P. Gough, L.C. Ehri, and R. Treiman (Eds.), *Reading acquisition* (pp. 65-106). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ventura, P., Kolinsky, R., Fernandes, S., Querido, L., & Morais, J. (2007). Lexical restructuring in the absence of literacy. *Cognition, 105*, 334-361.
- Wagenveld, B., Segers, P.C.J., Alphen, P.M. van & Verhoeven, L.T.W. (2013). The role of lexical representations and phonological overlap in rhyme judgments of beginning, intermediate and advanced readers. *Learning and Individual Differences, 23*, 64-71.
- Wagner, R., Torgeson, J. and Rashotte, C. (1994). Development of reading-related phonological processing abilities: new evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology, 30*, 73–87.
- Walley, A. C. (1987). Young children’s detections of word-initial and –final mispronunciations in constrained and unconstrained contexts. *Cognitive Development, 2*, 145-167.
- Walley, A. C., & Flege, J. E. (1999). Effect of lexical status on children’s and adults’ perception of native and nonnative vowels. *Journal of Phonetics, 27*, 307–332.
- Walley, A.C. & Metsala, J.L. (1990). The growth of lexical constraints on spoken word recognition. *Perception & Psychophysics, 47*, 267-280.
- Walley, A.C. & Metsala, J.L. (1992). Young children’s perception of spoken words varying in age-of-acquisition. *Memory & Cognition, 20*, 171-182.
- Walley, A. C., Metsala, J. L., & Garlock, V. M. (2003). Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. *Reading and Writing, 16*, 5–20.
- Werker, J.F. & Curtin, S. (2005). PRIMIR: A Developmental Framework of Infant Speech Processing. *Language Learning and Development, 1*, 197-234.

- Werker, J.F., Fennell, C.T., Corcoran, K.M., & Stager, C.L. (2002). Infants' ability to learn phonetically similar words: effects of age and vocabulary. *Infancy*, 3, 1–30.
- Westermann, G. & Reck Miranda, E. (2004). A new model of sensorimotor coupling in the development of speech. *Brain and Language*, 89, 393-400.
- Wicklegren, W. A. (1969). Context-sensitive coding, associative memory, and serial order (in speech) behavior. *Psychological Review*, 76, 1-15.
- Wilson, A. M. & Lesaux, N. K. (2001). Persistence of phonological processing deficits in college students with dyslexia who have age-appropriate reading skills. *Journal of Learning Disabilities*, 34, 394–400.
- Williams, R. J. & Zipser, D. (1995). Gradient-based learning algorithms for recurrent networks and their computational complexity. In Y. Chauvin and D. E. Rumelhart (Eds.) *Back-propagation: Theory, Architectures and Applications* (pp.433-482). Hillsdale, NJ: Erlbaum.
- Wolf, M., O'Rourke, G. A., Gidney, C., Lovett, M., Cirino, P., & Morris, R. (2002). The second deficit: an investigation of the independence of phonological and naming-speed deficits in developmental dyslexia. *Reading and Writing*, 15, 43–72.
- Wrench, A., and Hardcastle, W. (2000). *A multichannel articulatory speech database and its application for automatic speech recognition*, Paper presented at the Proceedings of the 5th Seminar on Speech Production.
- Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review*, 5, 683-689.
- Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, 131, 3-29.

Appendices

Appendix 1: Items for the mispronunciation conflict task. Global similarity distances between the items and the stimulus are shown in brackets.

	Target word	Two phoneme response	One phoneme response
Training item ^a	cat	<i>lat</i>	<i>lort</i>
Rime level items	pig	<i>nig</i> (5.40)	<i>teg</i> (5.32)
	kiss	<i>wiss</i> (5.2)	<i>tais</i> (5.33)
	bed	<i>yed</i> (5)	<i>vid</i> (5.02)
	pen	<i>len</i> (5)	<i>tain</i> (5.1)
	tail	<i>yail</i> (5.2)	<i>pell</i> (5.1)
	cage	<i>mage</i> (5.2)	<i>tedge</i> (5.3)
	bin	<i>yin</i> (5)	<i>ven</i> (5.02)
	fish	<i>nish</i> (4.8)	<i>saish</i> (4.73)
Phoneme level items	wave	<i>waish</i> (5)	<i>wib</i> (4.63)
	bib	<i>bik</i> (4.7)	<i>baiv</i> (4.63)
	neck	<i>nem</i> (5.2)	<i>nape</i> (5.5)
	net	<i>nesh</i> (5.1)	<i>nake</i> (5.5)
	chip	<i>chidge</i> (5)	<i>chet</i> (5.32)
	cave	<i>kaitch</i> (5.3)	<i>keb</i> (4.8)
	gate	<i>gaish</i> (5.1)	<i>gip</i> (4.93)
	shape	<i>shain</i> (5.4)	<i>shet</i> (5.1)

^aThe table headings do not apply to the training item. For the training item the mispronunciations consisted of one pseudoword sharing two phonemes with the target and one pseudoword sharing one phoneme with the target which was not globally matched.

Appendix 2: Items for the mispronunciation reconstruction task. Global similarity distances between the items and the stimulus are shown in brackets.

	Mispron- unciation	Two phoneme response	One phoneme globally matched response	One phoneme unmatched response	Unrelated word
Training item ^b	<i>handwich</i>	sandwich	helicopter	caterpillar	bus
Rime level items	<i>hain</i>	rain (5.00)	pin (5.23)	bone (10.06)	tap (14.11)
	<i>pell</i>	shell (5.00)	tail (5.1)	wheel (8.75)	bone (14.22)
	<i>tais</i>	race (5.20)	kiss (5.33)	juice (10.08)	pig (9.33)
	<i>ven</i>	ten (4.60)	bin (5.02)	man (7.75)	mud (14.22)
	<i>sis</i>	kiss (4.8)	face (4.73)	bus (8.62)	pan (14.25)
	<i>tain</i>	rain (5.20)	pin (4.93)	phone (10.16)	soap (15.56)
Phoneme level items	<i>nake</i>	name (5.20)	net (5.5)	nurse (9.8)	shed (11.8)
	<i>shet</i>	shell (4.70)	shape (5.1)	sheep (6.85)	man (13.35)
	<i>kaif</i>	cage (4.80)	kiss (4.73)	cap (8.81)	horse (11.86)
	<i>nape</i>	nail (5.00)	net (5.1)	nose (10.36)	hill (11.83)
	<i>tet</i>	ten (4.70)	tape (5.1)	teeth (7.65)	sun (14.42)
	<i>wib</i>	witch (4.80)	wave (4.63)	wall (9.71)	phone (15.11)

^bThe table headings do not apply to the training item. For the training item the target word differed by only one phoneme and the distracter words were chosen to be phonologically distant from the stimulus.

Appendix 3: Items for the pseudoword similarity task. Global similarity distances between the items and the stimulus are shown in brackets.

	Standard	Two phoneme response	One phoneme response
Training item ^c	Gary	Barry	Tom
Rime level items	<i>teyb</i>	<i>yeyb</i> (5.2)	<i>peb</i> (5.1)
	<i>bik</i>	<i>yik</i> (5)	<i>vek</i> (5.02)
	<i>fid</i>	<i>nid</i> (4.8)	<i>seyd</i> (4.73)
	<i>pim</i>	<i>nim</i> (5.4)	<i>tem</i> (5.32)
	<i>teych</i>	<i>yeych</i> (5.2)	<i>pech</i> (5.1)
	<i>beb</i>	<i>yeb</i> (5)	<i>vib</i> (5.02)
	<i>kib</i>	<i>wib</i> (5.2)	<i>teyb</i> (5.33)
	<i>pif</i>	<i>nif</i> (5.4)	<i>tef</i> (5.32)
Phoneme level items	<i>veyt</i>	<i>veysh</i> (5.1)	<i>vip</i> (4.93)
	<i>feb</i>	<i>fesh</i> (5.1)	<i>fiv</i> (5.02)
	<i>teyz</i>	<i>teyn</i> (4.9)	<i>tiss</i> (5.33)
	<i>leyp</i>	<i>leydge</i> (5)	<i>lek</i> (5.5)
	<i>bis</i>	<i>bim</i> (4.7)	<i>beyf</i> (4.73)
	<i>ses</i>	<i>sen</i> (5)	<i>sif</i> (5.12)
	<i>jeyt</i>	<i>jeysh</i> (5.1)	<i>jep</i> (5.1)
	<i>zeyp</i>	<i>zeyn</i> (5.4)	<i>zet</i> (5.1)

^cFor the training item one of the choices shared the same rime with the standard, the other was phonologically unrelated to the target.

Appendix 4: Items for the incomplete word task. Global similarity distances between the items and the stimulus are shown in brackets.

	Stimulus	Target word	Distracter 1	Distracter 2	Distracter 4
Training item ^d	s	sock	apple	brush	mouse
Single consonant rime level items	c	cup	pan	bag	mouse
	f	fish	van	zip	net
	s	sun	fish	van	bed
	b	boat	pig	tap	hat
	c	cat	gun	duck	feet
	g	gate	coat	pan	sheep
	m	mouse	net	pen	chin
	b	boot	girl	teeth	sheep
Consonant cluster phoneme level items	pr	pram	plate	coat	leaf
	br	bridge	black	pig	cheese
	sp	spade	smoke	zip	bike
	fl	flag	frog	van	badge
	cl	cloud	crown	tap	mud
	sk	skirt	spoon	house	moon
	gl	glove	grass	duck	feet
	cr	crab	clock	goat	van

^dThe table headings do not apply for the training item. For this item the distracters were chosen to all have phonetically distant onsets from the stimulus.

Appendix 5: Items for the implicit and explicit rhyme tasks. Global similarity distances between the items and the stimulus are shown in brackets.

	Stimulus ^e	Rhyming target	Two phoneme distracter	One phoneme distracter	Unrelated distracter
Training item	sat, bat, mat	cat	bike	wall	horse
	same, came, game	name (4.8)	gate (4.7)	gun (8.45)	bowl (14.76)
	tall, call, hall	ball (4.4)	hill (4.91)	shell (9.77)	sun (14.19)
	fed, bed, head	red (5)	hen (4.8)	hat (7.95)	nose (14.52)
	can, ran, pan	man (5)	pin (4.55)	moon (10.8)	beach (14.32)
	pick, lick, sick	chick (4.2)	sack (4.55)	book (8.62)	fish (7.5)
	coat, goat, note	boat (4.6)	nose (4.7)	nail (10.36)	sheep (14.3)
	peek, week, cheek	beak (4.8)	cheese (4.8)	chip (6.5)	net (12.75)
	fun, done, run	sun (5)	rain (4.55)	pan (10.12)	coat (14.12)

^eFor the explicit task the stimulus is the last word only, e.g. ‘Which one rhymes with mat?’

Appendix 6: Items for the blending and phoneme isolation tasks

	Stimulus ^f	Target	Shared rime	Shared onset	Unrelated
Training items	d-og + d-o-g	dog	bin	hat	leaf
Rime level items	f-ace	face	race	fish	pen
	m-ouse	mouse	house	moon	bat
	ch-in	chin	pin	cheese	leg
	r-ope	rope	soap	rain	bag
	ph-one	phone	bone	feet	bed
	b-in	bin	pin	bag	feet
	t-en	ten	pen	teeth	face
	c-at	cat	bat	can	bone
Phoneme level items	sh-e-d	shed	bed	sheep	boat
	b-al-l	ball	wall	bin	sheep
	h-ea-d	head	red	hat	cheese
	p-e-n	pen	men	pool	house
	l-e-g	leg	peg	leaf	mouse
	n-ur-se	nurse	purse	net	beak
	h-a-t	hat	cat	hen	ball
	p-a-n	pan	man	peg	coat

^fFor the phoneme isolation task the stimulus is just the picture of the target word.

Appendix 7: Items for the letter-sound knowledge task.

Single consonants		Digraphs
1. S s	14. U u	26. ch
2. A a	15. R r	27. sh
3. T t	16. H h	28. th
4. P p	17. B b	29. ng
5. I i	18. F f	30. ai
6. N n	19. L l	31. ee
7. M m	20. J j	32. igh
8. D d	21. V v	33. oa
9. G g	22. w w	34. oo
10. C c	23. X x	
11. O o	24. Y y	
12. K k	25. Z z	
13. E e		

Appendix 8: Matching statistics for tasks where the two and one phoneme choices were matched for global similarity. Standard deviations are shown in brackets.

Task	Mean dissimilarity between the stimulus and the response		Absolute dissimilarity difference		Paired sample test statistic
	two phoneme response	one phoneme response	Mean	Maximum	
Mispron. conflict and decision	5.10 (0.048)	5.09(0.075)	0.20(0.15)	0.50	t(15)=0.068, p=.95
Mispron. Recon.	4.92(0.21)	5.04(0.26)	0.22(0.13)	0.42	t(11)=-1.85, p=.091
Pseudo-word similarity	5.09(0.20)	5.11(0.21)	0.14(0.15)	0.50	t(15)=-0.33, p=.75
Implicit and explicit rhyme	4.13(1.69)	4.11(1.67)	0.26(0.21)	0.51	t(7)=0.14, p=.89

Appendix 9: Matching statistics for frequency

Task	Mean frequency (standard deviation in brackets)				Repeated measures ANOVA test statistic
	Target item	Distracter 1	Distracter 2	Distracter 3	
Mispronunciation reconstruction	2.84(0.70)	2.76(0.61)	2.96(0.46)	2.80(0.65)	$F(3,33) < 1, p = .87$
Incomplete word	2.83(0.52)	3.04(0.47)	2.85(0.72)	2.85(0.52)	$F(3,45) < 1, p = .69$
Implicit and explicit rhyme	3.07(1.18)	2.61(0.56)	2.65(0.66)	2.84(0.26)	$F(3,21) < 1, p = .57$
Blending	3.12(0.64)	3.22(0.66)	3.04(0.79)	3.10(0.77)	$F(3,45) < 1, p = .91$

Appendix 10: Matching statistics for phonotactic probability: positional segment average

Task	Mean positional segment average				Repeated measures ANOVA test statistic
	Target item	Distracter 1	Distracter 2	Distracter 3	
Mispron. conflict and decision	0.053	0.048	0.053	N/A	F(2,30)<1, p=.64
Mispron. reconstruction	0.059 (0.014)	0.064 (0.022)	0.053 (0.019)	0.069 (0.015)	F(1.43,15.74)=2.01, p=.13 ^g
Pseudoword similarity	0.055 (0.018)	0.046 (0.022)	0.053 (0.013)	N/A	F(2,30)=1.89, p=.17
Incomplete word	0.055 (0.016)	0.057 (0.011)	0.058 (0.013)	0.051 (0.015)	F(3,45)<1, p=.48
Implicit and explicit rhyme	0.064 (0.017)	0.063 (0.024)	0.061 (0.016)	0.054 (0.021)	F(3,21)<1, p=.68
Blending	0.063 (0.017)	0.068 (0.015)	0.057 (0.020)	0.057 (0.017)	F(2.17,32.53)=1.70, p=.20 ^g

^g With Greenhouse-Geisser correction for violation of assumption of sphericity.

Appendix 11: Matching statistics for phonotactic probability: biphone average

Task	Mean biphone average				Repeated measures ANOVA test statistic
	Target	Distracter	Distracter	Distracter	
	item	1	2	3	
Mispron. conflict and decision	0.0043	0.0036	0.0036	N/A	F(2,30)<1, p=.73
Mispron. reconstruction	0.0054 (0.0022)	0.0050 (0.0026)	0.0044 (0.0030)	0.0061 (0.0035)	F(1.69,18.53)=1, p=.51 ^h
Pseudoword similarity	0.0038 (0.0022)	0.0031 (0.0025)	0.0032 (0.0014)	N/A	F(2,30)=0.87, p=.43
Incomplete word	0.0049 (0.0029)	0.0053 (0.0026)	0.0051 (0.0029)	0.0039 (0.0024)	F(3,45)<1, p=.45
Implicit and explicit rhyme	0.0060 (0.0034)	0.0056 (0.0033)	0.0059 (0.0033)	0.0045 (0.0028)	F(3,21)<1, p=.78
Blending	0.0057 (0.0031)	0.0062 (0.0030)	0.0048 (0.0035)	0.0040 (0.0021)	F(2.07,30.97)=2.07, p=.14 ^h

^h With Greenhouse-Geisser correction for violation of assumption of sphericity

Appendix 12: Words used to train and test the neural network model

this	juice	his	dog	that	use
was	made	job	bog	much	big
this	had	mean	dig	yet	like
safe	mean	same	big	berg	could
those	have	thing	goat	pill	big
Jane	than	bought	not	John	fawn
than	nice	was	his	well	rich
yell	can	was	but	time	should
than	some	with	would	men	shoes
will	was	some	fog	much	sing
will	long	not	them	his	theme
can	was	that	time	food	shell
roll	coat	does	was	Jim	shock
wall	much	cows	with	Pam	that
home	can	this	full	why	was
can	get	light	moon	such	verse
use	will	lot	shone	gas	look
those	had	live	that	rug	cheap
love	third	church	night	this	gab
young	juice	with	Tom	Joyce	lone
rise	much	hot	chose	his	shone
did	his	have	watch	church	site
dad	was	can	chip	sun	this
beg	with	like	date	with	thing
that	sought	roof	sauce	right	cheese
one	his	chop	that	big	wrap
pick	book	wood	big	house	red
peck	bought	much	dog	nine	with
get	him	need	rag	like	sauce
cat	use	line	doll	long	thick
keep	with	then	Tod	hot	was
good	took	fill	top	young	Dutch
pop	down	not	his	mouse	shows
nice	would	have	bike	both	goose
young	this	pop	male	teeth	with
was	was	but	teeth	took	made
keep	not	look	not	word	his
with	nan	was	you'll	fish	him
can	will	red	get	leap	bob
that	house	rose	back	lake	that
was	was	his	than	Tim	not
caught	hit	wall	put	week	time
red	does	work	guss	can	those
mum	use	was	pine	roll	not
we'll	make	tag	his	one	
have	could	with	walk	will	
lack	rich	beach	rob	down	
heat	thick	tooth	made	jam	
had	tube	come	bob	we'll	
met	would	tooth	both	serve	
judge	did	fell	with	talk	
fail	would	birth	did	big	

will	file	jeff	shaun	can
got	serve	thought	catch	have
rain	this	rich	that	fish
gas	use	long	big	was
did	mine	run	goose	big
was	would	pays	have	top
fill	make	this	feet	hat
that	made	these	was	laugh
with	wool	take	take	sing
foam	young	this	was	ride
will	should	should	his	but
date	with	zoos	big	much
cheque	much	have	meet	with
deal	get	rock	this	fudge
with	get	roll	both	goes
good	not	has	have	well
should	juice	not	that	with
big	rob	these	bought	move
need	sat	shoes	both	lodge
with	geese	than	yacht	calf
have	wall	shape	home	this
can	with	like	will	has
save	jeff	cheese	this	team
tell	gave	but	dish	coach
cut	them	made	was	has
love	let	cheap	will	watch
live	them	run	not	thin
guess	set	time	down	dime
was	news	cheese	soon	with
was	web	with	save	those
call	sun	will	will	teach
dime	need	that	this	chain
live	put	not	ship	wash
hung	worm	hood	hull	light
have	hook	jeep	might	cash
his	have	was	night	will
was	with	hot	loss	top
hot	come	sun	with	have
sun	with	give	his	have
learn	put	one	will	moth
his	good	those	shoes	path
good	keep	will	bob	cab
back	tongue	them	room	but
bowl	could	five	that	him
bowl	cook	would	noise	tube
was	mum	need	right	should