



**Manchester
Metropolitan
University**

Ainsworth, Stephanie and Welbourne, Stephen and Woollams, Anna and Hesketh, Anne (2022) Global versus phonemic similarity : evidence in support of multi-level representation. *Cognition*, 225. p. 105138. ISSN 0010-0277

Downloaded from: <https://e-space.mmu.ac.uk/629645/>

Version: Published Version

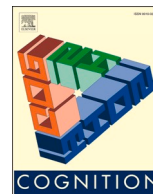
Publisher: Elsevier

DOI: <https://doi.org/10.1016/j.cognition.2022.105138>

Usage rights: Creative Commons: Attribution 4.0

Please cite the published version

<https://e-space.mmu.ac.uk>



Global versus phonemic similarity: Evidence in support of multi-level representation

Steph Ainsworth^{a,*}, Stephen Welbourne^b, Anna Woollams^b, Anne Hesketh^b

^a Faculty of Education, Manchester Metropolitan University, 53 Bonsall Street, Manchester M15 6GX, United Kingdom

^b School of Psychological Sciences, University of Manchester, M13 9PL, United Kingdom

ARTICLE INFO

Keywords:

Phonological representation
Lexical restructuring
Global similarity
Phoneme
Multi-level representation
Exemplar theory

ABSTRACT

There is long-standing debate about the extent to which children cognitively 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 within a cross-sectional ($N = 90$, aged 3;2 to 5;7) and longitudinal sample ($N = 23$, aged 3;2 to 5;1). The results show that children's sensitivity to phonemes increases over the first two years of school but does not reach adult levels. The findings indicate that global similarity relations remain important throughout development and support the idea of multi-level representation.

1. Introduction

A common point of contention across theoretical accounts of phonological development is the extent to which children represent phonological forms segmentally, (e.g., in terms of phonemes or onsets and rimes), or globally in terms of nonanalytic wholes (Ainsworth, Welbourne, Woollams, & Hesketh, 2019; Hallé & Cristia, 2012; Metsala & Walley, 1998; Saffran & Graf Estes, 2006; Thiessen, 2007; Vihman, 2017). Within the lexical restructuring model it is argued that children's phonological representations are initially holistic: based on the overall 'sounds-likeness' of words or on one particularly salient feature (Metsala & Walley, 1998). As children's vocabularies grow, their representations are proposed to become gradually segmented over childhood into onset-rime and then phonemic form. The idea of children's representations being initially based on global properties is supported by studies which show qualitative differences between the way that adults and children process words (Carroll & Myers, 2011; Cole & Perfetti, 1980; Treiman & Breaux, 1982; Walley, 1987; Walley & Metsala, 1990) and is consistent with observational studies of early speech production (Ferguson & Farwell, 1975; Vihman, 2017; Vihman & Croft, 2007) which suggest infants initially rely on word level templates that become increasingly segmental over development (Hallé & Cristia, 2012; Vihman & Croft, 2007).

In apparent contradiction with the evidence for early global representations, several studies have suggested that phonological

representations are fully specified from infancy (Bailey & Plunkett, 2002; Ballem & Plunkett, 2005; Swingley & Aslin, 2000; Swingley & Aslin, 2002; Swingley, 2009) based on infants' ability to detect minimal mispronunciations from as early as 14 months (see Ramon & Bosch, 2014 for a review). However, in her critical review of the experimental evidence, Vihman (2014) argues that this apparent contradiction might be resolved if we interpret the mixed experimental findings as evidence that:

'[r]epresentation is variable under differing conditions, with word production the most demanding and thus the most likely to reflect incomplete (or 'holistic') recall of the adult target form, particularly as regards unaccented syllables, voicing differences or codas (p. 210).

Another way to reconcile accounts of early specificity on the one hand and gradual emergence of phonemic representation on the other, is to reframe our conceptualisation of 'lexical restructuring' to mean the sharpening up of phonological categories across the lexicon as a whole rather than a qualitative change in the structure of individual representations (McKean, Letts, & Howard, 2013; Swingley, 2009). The PRIMIR (Processing Rich Information from Multidimensional Interactive Representations) framework (Werker & Curtin, 2005) is consistent with both of these framings, proposing that access to phonetic detail is task dependent and that adult-like phonemic categories are gradually honed over development (Werker & Curtin, 2005).

* Corresponding author.

E-mail addresses: s.ainsworth@mmu.ac.uk (S. Ainsworth), anna.woollams@manchester.ac.uk (A. Woollams).

PRIMIR proposes simultaneous representation across three multidimensional planes: the Perceptual, Word Form and Phoneme plane. 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 Phoneme plane. In this way, PRIMIR provides a candidate model for lexical restructuring that allows both early specificity (in the form of phonetically rich exemplars stored within the exemplars of the Word Form plane) and increasingly segmental representation (in the form of emerging phoneme-like categories in the Phoneme plane) to exist in parallel.

Although Werker and Curtin (2005) do not refer to the idea of global similarity directly, we suggest that this concept might be applied to the clustering of word level exemplars in terms of overall phonetic similarity within the Word Form plane. Following Carroll and Snowling (2001), we define global similarity as the overall sound similarity between whole words. Two words may be globally similar (sound alike) even if they do not share any phonemes, e.g. beach and dish:

This is because the phonemes within the words do share some features: the initial consonants are voiced stops, the vowels are close front vowels, and the final consonants contain post-velar frication (Carroll & Snowling, p. 328).

Within PRIMIR the gradual emergence of phonemes does not result from the restructuring of global 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. The majority of work has focussed instead on the emergence of segmental sensitivity (e.g. Ainsworth et al., 2019; Caudrelier et al., 2019; Foy & Mann, 2009; van den Bunt et al., 2018; Ventura, Kolinsky, Fernandes, Querido, & Morais, 2007). The current study manipulates global similarity and the number of shared phonemes within a mispronunciation-reconstruction task to allow measurement of both types of sensitivity over the first two years of school (age range 3;2 to 5;7).

2. Method

2.1. Participants

We present cross-sectional data from 90 children (46 boys, 44 girls), grouped according to age and school class: 'younger nursery' ($n = 24$, aged 3;2 to 3;10); 'older nursery' ($n = 24$, aged 4;0 to 4;5); 'younger reception' ($n = 22$, aged 4;0 to 4;7) and 'older reception' ($n = 20$, aged 4;7 to 5;7).¹ The younger nursery and younger reception groups were tested in the autumn term; the older nursery and older reception groups were tested in the late spring/summer terms. We chose these groupings to capture performance at different stages in children's developmental and educational journeys. Although there was substantial overlap

¹ In England, most children enter school in nursery (prekindergarten) classes aged 3 to 4, but compulsory education begins with entry to reception (kindergarten) classes, aged 4 to 5 years. Phonological awareness activities begin in nursery classes, following a systematic programme up until year 2 (aged 6 to 7 years).

between the ages of the older nursery and younger reception groups, they had received differing amounts of literacy instruction. The younger nursery group were tested over three additional time points roughly 5 months apart. One child from this group withdrew from the study after the second time point; their data were excluded from the longitudinal analyses, but included within the cross-sectional analyses. 74 adult undergraduate students were also included as a comparison group of literate adults.

The data were collected within a broader project (Ainsworth et al., 2019), where children and adults were tested with 4 measures of segmental sensitivity: *mispronunciation reconstruction* (the focus of this paper), *mispronunciation conflict* (child decides which of two mispronunciations sounds the most like the target), *pseudoword similarity* (child chooses which of two pseudowords sounds the most like the target pseudoword) and *initial sound* (child picks the picture corresponding to a spoken onset). These measures, which did not require any explicit awareness of the sounds in words were contrasted with 3 explicit segmental analysis measures (blending, phoneme isolation and rhyme). The blending task required children to listen to the pre-recorded voice of a robot say either an onset and a rime (e.g. *t-en*) or three individual phonemes (e.g. *t-e-n*); the phoneme isolation task asked children to say the sounds in a CVC word (e.g. *c-a-t*) when shown the corresponding picture; and the rhyme task asked children to choose the picture that rhymes with the word spoken by a puppet.

Measures of vocabulary and letter-sound knowledge were also taken. Expressive and receptive vocabulary were measured using the Renfrew Word Finding Vocabulary Test (Renfrew, 1997) and the British Picture Vocabulary Scale (Dunn et al., 2009) respectively; for the letter-sound knowledge task children were shown a grapheme (a letter or group of letters, e.g. 'sh') and asked to say what sound it represented (35 graphemes were included). The results from this first paper showed that 'although explicit segmental analysis is related to letter-sound knowledge, tasks measuring implicit segmental sensitivity provide evidence of segmental phonology related to vocabulary growth and not mediated by orthography' (Ainsworth et al., 2019, p. 323). Within the current paper, we perform additional analyses on the data from the mispronunciation reconstruction task. This design of this task allows us (for the first time to the author's knowledge) to generate concurrent measures of sensitivity to both phonemes and global similarity. Details of the other measures are provided in detail in Ainsworth et al. (2019).

2.2. The mispronunciation-reconstruction task

Children heard a puppet mispronounce a CVC word (spoken live by the researcher) 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 (represented pictorially on cards). For each trial the child was presented with four response choices:

Two-phoneme response: a word sharing two phonemes with the stimulus (e.g. *hain* – rain)

- 1) 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. *hain* – pin).
- 2) 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. *hain* – bone).
- 3) Unrelated response: this word shares no phonemes with the stimulus and is also globally distant (e.g. *hain* – tap).

If a child chooses response 1 more often than response 2, 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 a child chooses response 3 more often than response 4 we can infer that they are sensitive to global similarity when phonemic similarity is held

constant. The task consisted of a training trial plus 12 trials (see Ainsworth et al., 2019 for list of stimuli). Corrective feedback was given for the training trial only.

For adults, the same list of stimuli was used, but the procedure was adapted for an adult audience. For adults the pictures appeared on a screen alongside pre-recorded auditory stimuli, presented using *E-Prime*.²

2.2.1. 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). 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 were matched listwise for frequency and phonotactic probability (see Ainsworth et al., 2019 for matching characteristics).

2.2.2. 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) 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 /e/ (2.22) plus the dissimilarity of /n/ and /d/ (4.8).

This metric provides an approximate value for how distant two stimuli are from one another in terms of global similarity. It is important to note that the metric does not take into account the effects of co-articulation and the way that words are perceived incrementally rather than all at once. It has been found, however, to be a useful and valid measure in previous studies (Byrne & Fielding-Barnsley, 1993; Carroll & Snowling, 2001; Wagensveld, Segers, van Alphen, & Verhoeven, 2013), which found that global similarity, operationalised in this way had a strong confounding influence on children's phonological similarity judgements.

3. Results

Data screening excluded one adult who had extreme outlier scores for three of the six tasks within the full battery (reported in Ainsworth et al., 2019). In all analyses, unless otherwise stated: only the child data (and not the adult data) were included; significance values are one-tailed; and error bars represent confidence intervals of 95%.

3.1. Cross-sectional results

A plot of response frequency by response type (Fig. 1) showed a steady developmental increase in the number of two-phoneme responses with adult data included for comparison.

In order to test the hypothesis that over the first two years of school (in England) children become increasingly sensitive to the number of shared phonemes over and above the global similarity between words, analyses taken from signal detection theory were conducted. Following Massaro (1989), we treated the proportion of two-phoneme responses as hits and the one-phoneme globally matched responses as false alarms, allowing us to calculate a measure of phoneme sensitivity, d' as the

absolute difference between the transformed hit and false alarm rates (Macmillan & Creelman, 2005). To avoid the problem of undefined z scores, for proportions of 0 and 1 (e.g. when children always or never choose the two-phoneme response), the d' values were calculated using average proportions of two-phoneme responses and one-phoneme globally matched responses calculated within each group (young nursery, old nursery, etc.). This pooling of data to address the problem of extreme scores within signal detection studies is commonly conducted (Macmillan & Kaplan, 1985; Sussman, 1993) with collapsed values of d' from averaged proportions providing 'a reliable, relatively unbiased way to estimate true average d' ' (Macmillan & Kaplan, 1985, p. 196).

Similarly a measure of global sensitivity was calculated for each group as the absolute difference between the z scores of the average proportion of one-phoneme globally matched and one-phoneme unmatched responses. These values of d' prime represent children's sensitivity to global similarity over and above the number of shared phonemes. The cross-sectional profiles of phoneme and global sensitivity are plotted in Fig. 2. For comparison, the adult values of d' representing phoneme and global sensitivity were calculated as 1.68 and 0.29 respectively.

Calculations of the G statistic (Gourevitch & Galanter, 1967) were made to assess whether the children's phoneme and global sensitivity were significantly above chance. The G statistic is a parametric measure of significance which can be used to compare group measures of d' (Sussman, 1993). While the young nursery children's sensitivity to phonemes over and above global similarity was not yet significantly above chance ($G = 1.28, p = .10$), the old nursery children were already performing above chance ($G = 2.55, p = .005$). None of the four groups demonstrated significant levels of sensitivity when taken individually, but when pooled into two larger groups, nursery and reception, the reception ($G = 2.06, p = .020$) but not the nursery ($G = 1.16, p = .12$) children demonstrated a level of global sensitivity that was significantly greater than you would expect by chance.

G statistics were also calculated to test whether the differences in sensitivity between the groups in Fig. 2 were significant. For phoneme sensitivity, while the differences between the consecutive groups of children was found to be non-significant ($p > .05$), there was a significant difference between the sensitivity of the young nursery children and the old reception children ($G = 3.04, p = .001$) and between the old reception children and the adults ($G = 6.93, p < .00005$). None of the between-group differences were found to be significant for global sensitivity ($p > .05$).

3.2. Longitudinal results

The response profile for the longitudinal sample shown in Fig. 3 shows a similar pattern to the cross-sectional data, with the number of two phoneme responses rising steadily over development.

Values of d' were calculated in the same way as for the cross-sectional data to reflect sensitivity to phonemes and global similarity respectively (Fig. 3). Calculation of the G statistic (Gourevitch & Galanter, 1967) showed a significant rise in phoneme sensitivity from time 1 to time 4 ($G = 4.44, p < .00005$) and from time 3 to time 4 ($G = 1.90, p = .029$). No significance difference between the time points was found for global sensitivity ($p > .05$).

3.3. Comparison of cross-sectional and longitudinal performance

While Figs. 2 and 4 show a similar shape for the trajectory of phoneme and global sensitivity, it is noted that the value of phoneme sensitivity achieved by the longitudinal group is significantly higher than for the cross sectional group ($G = 96, p < .00005$). To investigate if this was due to cohort differences, performance on letter-sound knowledge, vocabulary measures and phonological awareness measures was compared between each cross-sectional group and the corresponding longitudinal data (Table 1). The younger nursery scores, which

² The auditory stimuli for the adult task were pre-recorded by the same researcher who spoke them live during the child-version of the task.

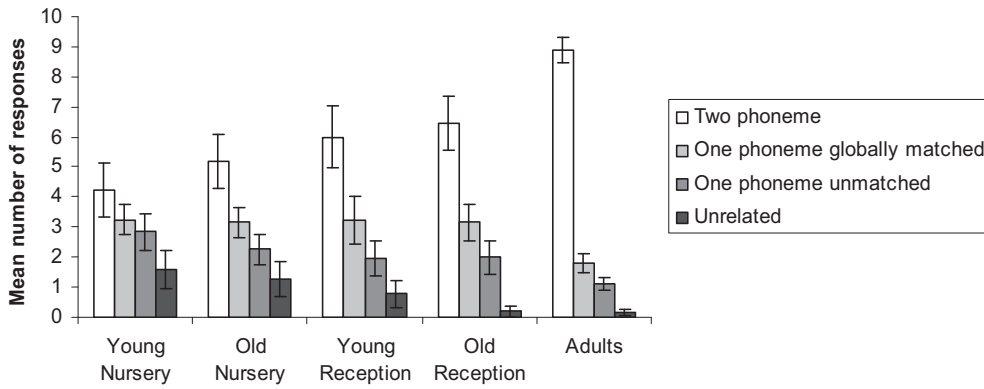


Fig. 1. Children's response profiles across four cross-sectional groups. Adult data included for comparison.

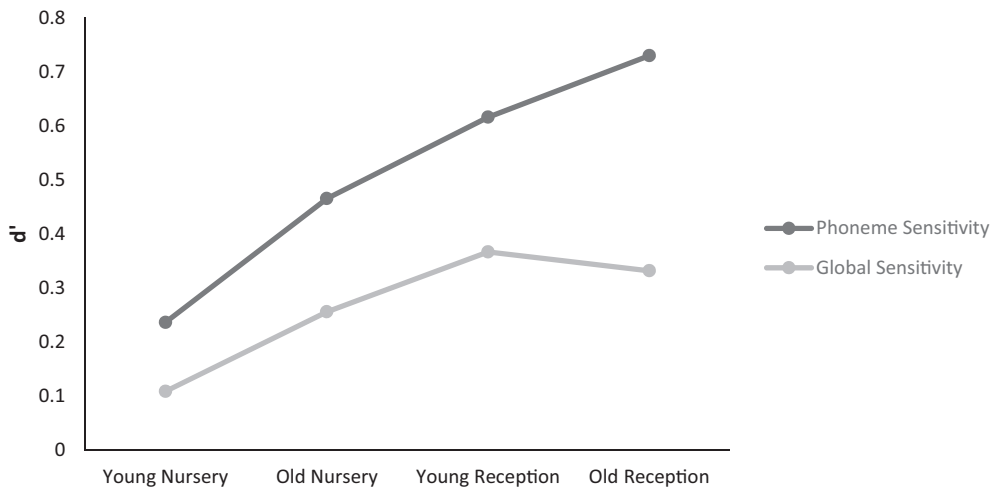


Fig. 2. Developmental plot of children's sensitivity to phonemes and global similarity from the cross-sectional data.

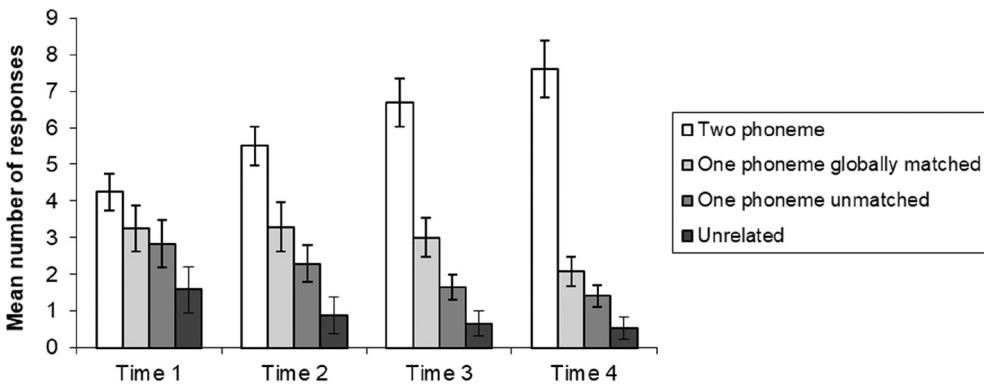


Fig. 3. Children's response profiles across four different time points.

are the same for both samples are included for completeness. The results show that the longitudinal sample performed significantly better than the cross-sectional sample on rhyme at time point 2, ($t(44) = 3.58, p = .0009$), and letter sound knowledge, ($t(43) = 2.26, p = .029$) blending ($t(43) = 2.28, p = .024$) and phoneme isolation performance ($t(43) = 2.35, p = .022$) at time point 3.

4. Discussion

The longitudinal and cross-sectional data provide convergent evidence of a developmental increase in phoneme sensitivity. Children's

sensitivity to phonemes was found to climb over their first two years of schooling (starting at age three in England), but not yet reaching adult levels. This observed rise in phoneme sensitivity is consistent with the lexical restructuring model (Metsala & Walley, 1998), the PRIMIR framework (Werker & Curtin, 2005), and other emergent theories of phonological development which predict phonemic representation to emerge gradually over development (e.g. Ventura et al., 2007; Ziegler & Goswami, 2005). The observed growth in phoneme sensitivity is also consistent with children's performance on other similarity judgment tasks which control for global similarity (Ainsworth et al., 2019).

Descriptively, Figs. 2 and 4 suggest that global sensitivity also

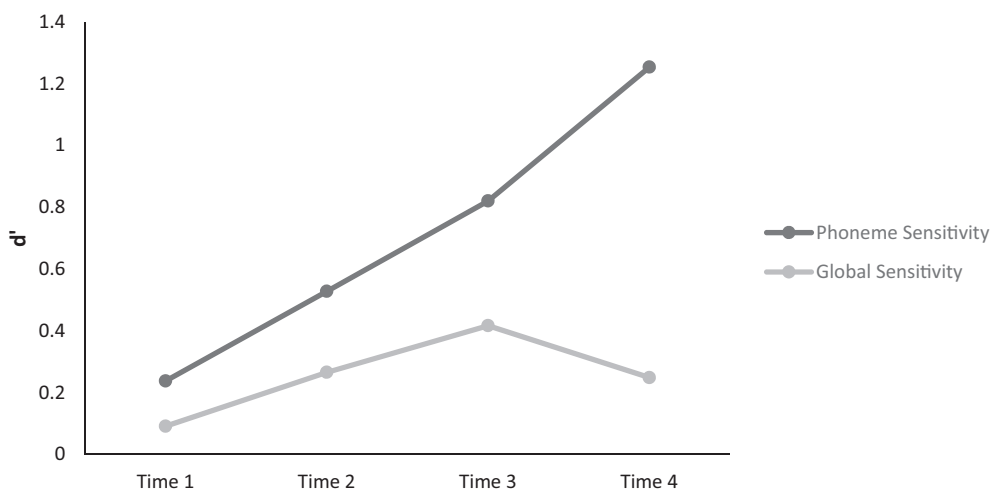


Fig. 4. Developmental plot of children's sensitivity to phonemes and global similarity from the longitudinal data.

Table 1

Means and standard deviations (in parentheses) comparing performance across cross-sectional and longitudinal samples.

Group	Letter-sound knowledge	Receptive vocabulary	Expressive vocabulary	Rhyme	Blending	Phoneme isolation
Younger nursery	2.57 (5.53)	47.17 (17.76)	17.85 (7.37)	33.15(19.07)	46.63(25.21)	13.41(26.04)
Older nursery	7.13 (8.77)	56.78 (17.64)	21.35 (6.29)	34.24(17.64)	64.13(19.60)	35.72(36.80)
Time 2	10.09 (8.85)	57.17 (12.68)	23.17 (6.55)	56.52(24.09)	71.47(21.39)	35.96(32.49)
t(44)	1.139	0.086	0.961	3.579	1.213	0.023
p	.26	.93	.34	.0009*	.23	.98
Younger reception	13.36 (9.78)	59.77 (13.32)	24.73 (4.28)	46.59(25.05)	75.03(22.67)	46.21(39.16)
Time 3	19.22 (7.48)	67.48 (10.91)	25.91 (7.31)	54.89(35.90)	88.80(17.60)	73.48(38.75)
t(43)	2.26	2.128	0.657	0.896	2.282	2.348
p	.029*	.039	.51	.38	.028*	.024*
Older reception	29.25 (3.37)	70.35 (11.90)	27.95 (5.81)	61.25(25.94)	82.5(10.65)	92.71(14.63)
Time 4	29.54 (5.40)	73.32 (11.31)	29.73 (6.20)	69.05(27.84)	89.29(20.65)	92.66(20.58)
t(41)	0.290	2.97	0.967	0.957	1.319	0.009
p	.84	.41	.34	.34	.19	.99

increases from the beginning of nursery to the beginning of the reception year before levelling off in the second half of reception; however, the changes in global sensitivity over time were not statistically significant. The fact that phoneme and global sensitivity appear to rise concurrently (descriptively) over the nursery year suggests that there is no 'trade off' associated with the rise in phoneme 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 phonetically rich exemplars being added to the Word Form plane as children gain more language experience. Adult performance suggests 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. Again, this is consistent with PRIMIR's idea of simultaneous levels of representation throughout development. Our results, therefore, support the notion of restructuring of the lexicon as a whole rather than of individual words being transformed from whole to parts representation (Mckean et al., 2013).

Given the failure of the observed rise in children's global similarity to reach significance, however, these conclusions remain speculative, and warrant further investigation using a larger sample size. Further research would also benefit from exploring alternative ways to avoid the problem of extreme scores on the task, allowing sensitivity to phonemes

and global similarity to be analysed at the individual level. While a number of corrections may be applied to extreme scores to allow individual analyses to take place each is associated with a trade off in terms of potentially biasing estimates of d' (Macmillan & Kaplan, 1985). A preferable solution involves redesigning tasks so that very high and very low levels of sensitivity are unlikely to occur (ibid). Consideration of how the mispronunciation reconstruction task might be amended in this way might be fruitful.

While the developmental profiles of the cross-sectional and longitudinal samples are broadly in line with one another, it is noted that the oldest children in the reception sample reached a lower level of phoneme sensitivity ($d' = 0.73$) than that achieved by the longitudinal sample by the end of the reception year ($d' = 1.25$). This suggests a shallower rise in phoneme sensitivity across the cross-sectional groups than between the longitudinal time points. This discrepancy might be due to cohort effects – children in the longitudinal sample happened to be more linguistically able on average than the children in the cross-sectional sample – or practice effects – where the very process of engaging with the task on multiple occasions improved levels of phoneme sensitivity. While the latter explanation is possible, it is important to note that there were several months between the four sessions.

Another possibility that needs to be considered is whether the observed rises in sensitivity are not true reflections of representational change, but rather reflect children simply getting better at attending to the task. This possibility is unlikely given that Ainsworth et al. (2019) showed the composite measure of segmental sensitivity which included the mispronunciation reconstruction task, calculated for the mispronunciation reconstruction task was predicted by vocabulary but not age

(for the cross-sectional data). This suggests that improved performance over development is due to representational changes, driven by lexical growth, rather than by general maturity.

One potential factor which might influence cohort effects is differences in educational experience, especially in relation to exposure to phonological awareness training and orthographic knowledge. The longitudinal sample performed significantly better than the cross-sectional sample on rhyme awareness at time point 2, and letter sound knowledge, blending and phoneme isolation performance at time point 3. As discussed in Ainsworth et al. (2019), it is possible that these higher scores reflect greater experience with letters and sounds (perhaps through cohort differences in the teaching of phonics), which, in turn, may have boosted phoneme sensitivity within the longitudinal cohort. This is in line with Ziegler and Goswami's (2005) proposition that acquisition of the alphabetic principle precipitates phonemic representation. On the other hand, the directionality might act in the other direction, with higher levels of phoneme sensitivity setting the stage for explicit phoneme awareness (Ventura et al., 2007).

Given the uncertainty in distinguishing between these interpretations, further work with a larger sample size is needed to get a clearer picture of the typical development of these two types of sensitivity. The potential for practice effects to improve sensitivity is of particular interest given the need for interventions which support the development of phoneme sensitivity at an age where children have not yet developed the metacognitive skills and letter knowledge required to complete traditional phonological awareness tasks (Ainsworth et al., 2019; Claessen, Heath, Fletcher, Hogben, & Leitão, S., 2009).

While we have interpreted our results as being most consistent with PRIMIR (Werker & Curtin, 2005) – a framework rooted in the speech perception literature – evidence for representation across multiple phonological units is also growing within the field of speech production (Tilsen, 2016; Vihman, 2017). For example, Vihman and Croft's development theory, which suggests that speech production units might operate in parallel across a range of grain sizes, has found recent support from an auditory-motor adaptation study with adults, showing parallel transfer at the word, syllable and phoneme level (Caudrelier, Schwartz, Perrier, Gerber, & Rochet-Capellan, 2018). While the debate around phonological development has historically been centred around establishing *what* the basic unit of speech production/perception is and whether this changes over development, it is perhaps more appropriate to ask *how* children might use the multiple levels of information available to them functionally during everyday perception and production events. This study contributes to this shift in focus by providing evidence for the continuing influence of global similarity alongside growing sensitivity to phonemes, which in turn, suggests that both children and adults use global information at the lexical level alongside phonemic information in similarity judgment tasks.

CRedit authorship contribution statement

Steph Ainsworth: Conceptualization, Methodology, Data curation, Writing – original draft. **Stephen Welbourne:** Supervision, Conceptualization, Methodology, Writing – review & editing. **Anna Woollams:** Supervision, Writing – review & editing. **Anne Hesketh:** Supervision, Conceptualization, Methodology, Writing – review & editing.

Declaration of Competing Interest

None.

Acknowledgements

We thank the children, parents, and teachers who were involved in this study. This work was supported by the Economic and Social Research Council (ES/J500094/1).

References

- Ainsworth, S., Welbourne, S., Woollams, A., & Hesketh, A. (2019). Contrasting explicit with implicit measures of children's representations: The case of segmental phonology. *Language Learning*, 69, 323–365. <https://doi.org/10.1111/lang.12334>
- Bailey, T. M., & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development*, 17(2), 1265–1282. [https://doi.org/10.1016/S0885-2014\(02\)00116-8](https://doi.org/10.1016/S0885-2014(02)00116-8)
- Ballem, K. D., & Plunkett, K. (2005). Phonological specificity in children at 1;2. *Journal of Child Language*, 32(1), 159–173. <https://doi.org/10.1017/S0305000904006567>
- van den Bunt, M. R., Groen, M. A., Frost, S., Lau, A., Preston, J. L., Gracco, V. L., & V. L., Pugh, K. R., & Verhoeven, L. T. W. (2018). Sensorimotor control of speech and Children's Reading ability. *Scientific Studies of Reading*, 22(6), 503–516. <https://doi.org/10.1080/10888438.2018.1491583>
- Byrne, B., & 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(1), 104–111.
- Carroll, J. M., & Myers, J. M. (2011). Spoken word classification in children and adults. *Journal of Speech, Language, and Hearing Research*, 54(1), 127–147. [https://doi.org/10.1044/1092-4388\(2010\)08-0148](https://doi.org/10.1044/1092-4388(2010)08-0148)
- Carroll, J. M., & Snowling, M. J. (2001). The effects of global similarity between stimuli on performance on rime and alliteration tasks. *Applied Psycholinguistics*, 22(3), 327–342. <https://doi.org/10.1017/S0142716401003034>
- Caudrelier, T., Ménard, L., Perrier, P., Schwartz, J.-L., Gerber, S., Vidou, C., & Rochet-Capellan, A. (2019). Transfer of sensorimotor learning reveals phoneme representations in preliterate children. *Cognition*, 192, Article 103973. <https://doi.org/10.1016/j.cognition.2019.05.010>
- Caudrelier, T., Schwartz, J., Perrier, P., Gerber, S., & Rochet-Capellan, A. (2018). Transfer of learning: What does it tell us about speech production units? *Journal of Speech Language and Hearing Research*, 61(7), 1613–1625. <https://doi.org/10.1044/2018.JSLHR-S-17-0130>
- 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(2), 121–124. <https://doi.org/10.1080/13682820801966317>
- 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. [https://doi.org/10.1016/S0022-5371\(80\)90239-X](https://doi.org/10.1016/S0022-5371(80)90239-X)
- 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.
- Ferguson, C. A., & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language*, 51(2), 419–439. <https://doi.org/10.2307/412864>
- 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(2), 339–361. <https://doi.org/10.1017/S0142716409090146>
- Gourevitch, V., & Galanter, E. (1967). A significance test for one parameter isosensitivity functions. *Psychometrika*, 32(1), 25–33.
- Hallé, P., & Cristia, A. (2012). Global and detailed speech representations in early language acquisition. In I. S. Fuchs, & P. Perrier (Eds.), *Speech planning and dynamics* (pp. 5–31). Berlin: Peter Lang Verlag. <https://doi.org/10.3726/978-3-653-01438-9>
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, N.J: Lawrence Erlbaum Associates.
- Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin*, 98(1), 185–199.
- Massaro, D. W. (1989). Testing between the TRACE model and the fuzzy logical model of speech perception. *Cognitive Psychology*, 21(3), 398–421.
- Mckean, C., Letts, C., & Howard, D. (2013). Functional reorganization in the developing lexicon: Separable and changing influences of lexical and phonological variables on children's fast-mapping. *Journal of Child Language*, 40(2), 307–335. <https://doi.org/10.1017/S0305000911000444>
- 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.
- Ofsted. (2014, May 12). Data dashboard. Retrieved from <http://dashboard.ofsted.gov.uk/>.
- Renfrew, C. (1997). *Word finding vocabulary test*. Milton Keynes: Speechmark.
- Saffran, J. R., & Graf Estes, K. (2006). Mapping sound to meaning: Connections between learning about sounds and learning about words. *Advances in Child Development and Behavior*, 34, 1–38. [https://doi.org/10.1016/S0065-2407\(06\)80003-0](https://doi.org/10.1016/S0065-2407(06)80003-0)
- Singh, S., & Woods, D. R. (1971). Perceptual structure of 12 American English vowels. *Journal of the Acoustical Society of America*, 49, 1861–1866. <https://doi.org/10.1121/1.1912592>
- Singh, S., Woods, D. R., & Becker, G. M. (1972). Perceptual structure of 22 prevocalic English consonants. *Journal of the Acoustical Society of America*, 52(6), 1698–1713. <https://doi.org/10.1121/1.1913304>
- 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.
- Sussman, J. E. (1993). Perception of formant transition cues to place of articulation in children with language impairments. *Journal of Speech and Hearing Research*, 36, 1286–1299.
- Swingle, D. (2009). Onsets and codas in 1.5-year-olds' word recognition. *Journal of Memory and Language*, 60(2), 252–269.

- Swingle, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, 76(2), 147–166.
- Swingle, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, 13(5), 480–484. <https://doi.org/10.1111/1467-9280.00485>
- Thiessen, E. D. (2007). The effect of distributional information on children's use of phonemic contrasts. *Journal of Memory and Language*, 22(3), 239–243. <https://doi.org/10.1016/j.jml.2006.07.002>
- Tilsen, S. (2016). Selection and coordination: The articulatory basis for the emergence of phonological structure. *Journal of Phonetics*, 55, 53–77. <https://doi.org/10.1016/j.wocn.2015.11.005>
- 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. <https://doi.org/10.1007/BF01067613>
- Ventura, P., Kolinsky, R., Fernandes, S., Querido, L., & Morais, J. (2007). Lexical restructuring in the absence of literacy. *Cognition*, 105(2), 334–361. <https://doi.org/10.1016/j.cognition.2006.10.002>
- Vihman, M. M. (2014). *Phonological development: The first two years* (2nd ed.). Oxford: Wiley.
- Vihman, M. M. (2017). Learning words and learning sounds: Advances in language development. *British Journal of Psychology*, 108(1), 1–27. <https://doi.org/10.1111/bjop.12207>
- Vihman, M. M., & Croft, W. (2007). Phonological development: Toward a 'radical' templatic phonology. *Linguistics*, 45(4), 683–725. <https://doi.org/10.1515/LING.2007.021>
- Wagensveld, B., Segers, E., van Alphen, P., & Verhoeven, L. (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. <https://doi.org/10.1016/j.lindif.2012.09.007>
- Walley, A. C. (1987). Young children's detections of word-initial and –final mispronunciations in constrained and unconstrained contexts. *Cognitive Development*, 2(2), 145–167. [https://doi.org/10.1016/S0885-2014\(87\)90110-9](https://doi.org/10.1016/S0885-2014(87)90110-9)
- Walley, A. C., & Metsala, J. L. (1990). The growth of lexical constraints on spoken word recognition. *Perception & Psychophysics*, 47, 267–280. <https://doi.org/10.3758/BF03205001>
- Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development*, 1(2), 197–234. <https://doi.org/10.1080/15475441.2005.9684216>
- Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, 131(1), 3–29. <https://doi.org/10.1037/0033-2909.131.1.3>