UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

The Development of Spolken Word Recognition: Experimental and Computational Studies

Permalink

https://escholarship.org/uc/item/64p3g17t

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 20(0)

Authors

Loucas, Tom
Marslen-Wilson, William D.

Publication Date

1998

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

The Development of Spoken Word Recognition: Experimental and Computational Studies

Tom Loucas (t.loucas@psyc.bbk.ac.uk)
Centre for Speech and Language, Birkbeck College, London WC1E 7HX

William D. Marslen-Wilson (w.marslen-wilson@mrc-apu.cam.ac.uk)
MRC Cognition and Brain Sciences Unit, Cambridge, CB2 2EF

Abstract

Children's spoken word recognition is little understood compared to our knowledge of the adult system. We present here a combined experimental and computational exploration of the development of lexical access. Three accounts of the way children represent lexical form (Full-Specification, Radical Underspecification and Gradual Segmentation) are rejected in favour of one which derives from a connectionist approach. It sheds light on the pattern of results from two experiments investigating the way children, aged 5- to 9-years-old, process regular and irregular variation in the surface form of speech, which suggested, whilst children's lexical representations are functionally underspecified from at least 5-years-old, they are only beginning to track the viability of regular phonological variation at 9-years-old. The late acquisition of phonological inference is accounted for in a connectionist model in terms of the sparseness of the information relevant to learning this structural relationship in language.

Introduction

The nature of spoken word recognition in adults is well elaborated. The same cannot be said for children and the developing system, where basic questions regarding lexical representation and processing remain unanswered. Central among these is: How do children represent lexical form? We address this question here using a combined experimental and computational approach to the way children develop the ability to deal with phonological variation in speech.

Phonologically regular variation has proven to be a great boon to elucidating the properties of the spoken word recognition in adults. Phonological processes operate in connected speech to phonetically alter surface forms (SFs), relative to lexically stored underlying representations (URs), in predicable ways. In English, for example, place assimilation affects word-final coronal segments, which acquire the place feature of word-initial noncoronal consonants of the following word. For example *boat* becomes [bəʊp] in the context *boat painter* and *chase* becomes [tʃeɪʃ] in the context *chase sheep*. Another example is fricative devoicing, which occurs when a word-final voiced fricative precedes a voiceless consonant. For example, *cheese* becomes [tʃi:s] in the context *cheese sandwich*.

This sort of variation is a problem for a spoken word recognition system which is highly intolerant to mismatch between sensory input and URs (Marslen-Wilson, 1993). Lahiri and Marslen-Wilson (1991), dissolve this dilemma by proposing that URs are highly abstract. Regular variation does not lead to mismatch because the dimensions on which SFs vary are not represented lexically. In their view, consistent with Radical Underspecification theory (Archangeli, 1988), only information which is unpredictable is included in URs.

However, a purely representational theory cannot wholly account for the way SFs are processed. Gaskell and Marslen-Wilson (1996) show that listeners are sensitive to the viability of assimilations. Adults will accept SFs at their assimilated off-sets, but they also track the phonological viability of SFs using, what Gaskell and Marslen-Wilson call, "phonological inference". Thus, the assimilated form "*leam*" will as act a prime for LEAN in the context "*...leam bacon...*" if probed at the offset of "*leam*" or at the onset of "*bacon*". But "*leam*" does not act as a prime the in the context "*...leam gammon...*" if probed at the onset of "*gammon*" because the [g] makes the assimilation unviable. These results are incompatible with Radical Underspecification which predicts there is no mismatch at any point during recognition because the place feature of a word-final coronal is not specified, and so there is no representational basis for any mismatch or viability effect.

Gaskell, Hare and Marslen-Wilson (1995) provide a computational account of abstractness and phonological inference. They trained a simple recurrent network (Elman 1990) to map between a continuous sequence of segments and current, previous and next segment, using a corpus of spoken English in which 0.5% of the words were place assimilated. The varying SFs were mapping to canonical URs. The trained network exhibited both representational abstractness—surface noncoronals were treated more variably than surface coronals—and the ability to do phonological inference—phonetic context was used to infer the underlying identity of SFs. Gaskell et al. described the model's representations as functionally underspecified, rather than radically underspecified, because their basis was an asymmetry in the mapping between variant SFs and URs, developed in response to the statistical properties of the training environment. The authors view both abstraction and inference as aspects of this one mapping process.

The scant attention paid by developmental psycholinguists to the issue of spoken word recognition in children is in marked contrast to an extensive literature on infant speech perception. A basic finding of this research is that by the end of the first year of life babies are able to perceptually analyse speech into the phonemic categories of their native language (Werker and Lalonde, 1988). This may well explain the lack of interest in lexical access in children, for it argues that they come to word learning with a pre-lexical phonological processor which allows them to phonemically analyse speech input. If young children are able to chunk the speech stream into phonemes, they could use this segmented input to build lexical representations which are then used as perceptual targets in lexical access. This implies that even young children's spoken word recognition should be adult-like in all its essentials. Consistent with this view, Gerken, Murphy & Aslin (1995) argue that lexical representations in children, as in adults, are segments consisting of phonetic-features—i.e. they are canonical, phonologically fully-specified representations.

Where psycholinguists have tended to assume unquestioningly that lexical representations are fully-specified, phonological theory has asked how much information is needed in the UR and whether there are changes in featural information during development. Gierut (1996) uses Radical Underspecification theory to argue that initial representations are only specified for universal distinctive properties and phonological acquisition is a process of adding marked nonredundant features.

A third, and sharply contrasting, view is that elaborated by Walley (1993), who questions the interpretation of infant speech perception as demonstrating an early ability to phonemically segment speech. She argues that URs develop over a prolonged period, lasting into middle childhood, from initially holistic representations, based on a word's overall acoustic shape with little internal structure, to phonemically segmented forms.

There is a fourth alternative. Gaskell et al. (1995) show it is possible to acquire representational abstractness and a sensitivity to the structural properties of language, expressed in phonological rules, such as place assimilation, in mapping between the surface form of speech and the lexicon using a connectionist learning mechanism. If children learning language are engaged in an equivalent process, then tracking the performance of a network as it learns this mapping should shed light on actual patterns of development. We initially test the three theories of representation outlined above and then compare the empirical data which results with a connectionist model of acquisition.

Experiment 1

Three contrasting theories of lexical representation in children—Full-Specification, Radical Underspecification, and Gradual Segmentation—have been identified. We exploited regular and irregular phonological variation to test these hypotheses in single word processing, using a speeded auditory lexical decision task.

English place assimilation and fricative devoicing were used to construct of three sets of 25 test stimuli. All test items were monosyllabic words which undergo assimilation[1]. In the Word condition, items were left in their canonical form. In the Assimilated condition they were produced as assimilated forms (e.g. *boat* > [bəʊp], *cheese* > [tʃiːs]). These are words in continuous speech, but nonwords out of an appropriate phonetic context. In the Nonword condition, items were randomly changed by a single feature on word-final segments (e.g. *dress* > [drɛf], *shut* > [ʃʌd]). This condition provided irregular phonetic variation to contrast with the regular variation of the Assimilated condition. The stimuli were matched for frequency, syntactic class, number of segments and the phonetic properties of onsets and codas.

Participants listened to these words one a time and made speeded lexical decisions to each. The hypotheses described above make different predictions for this task. If young children's URs are not yet fully segmented, they will not be sensitive to the single feature deviations used here. Both assimilated words and nonwords should be acceptable as real words. High error rates would be predicted for both and decision latencies should not differ. If the young children's representations are underspecified, they will show slower decision latencies and more errors for assimilated forms than nonwords because the assimilated forms will not mismatch with URs on their altered feature, whereas nonwords do. If children's representations are fully-specified they should reject assimilated items and nonwords equally quickly and show similar rates of errors in both conditions.

Walley (1993) suggests that there are changes in children's lexical representations across middle childhood, and so two groups of children were selected to span this period of development. Thirteen 5-year-olds (mean age = 5;10; range: 5;03 to 6;04), 12 8-year-olds (mean age = 8;07; range: 8;02 to 9;03), and 13 adults participated in Experiment 1.

Results and Discussion

Analyses of Covariance were used to analyse the reaction time data, with Item-Duration included as a covariate, because there were significant differences between the mean word durations between conditions. The mean decision latencies and item-durations are given in Table 1 and error rates in Table 2. An overall by-items ANCOVA of the 5-year-olds decision latencies showed a significant effect of Word-Type ($F2[2,68]=15.18$, $p<0.001$). Bonferroni corrected pairwise comparisons showed significant differences between Word and Assimilated items

---

[1] The words, all found in the productive vocabularies of 2- and 3-year-olds, were taken from the Wells (1981), Fletcher and Garman (1988) and Johnson (1986) corpora in the CHILDES database.

(F2[1,45]=24.28, p<0.001) and between Word and Nonword items (F2[1,45]=13.15, p=0.001). The difference between the critical Assimilated and Nonword conditions was not significant (F2[1,45]=1.70). There were no differences in error rates between conditions (F2[2,68]<1).

Table 1: Decision Latencies and Item-Durations for Experiment 1 in msec (standard errors for reaction times and item-durations are given in parentheses).

	Word	Assimilated	Nonword
Adults	778 (16)	803 (18)	780 (12)
8-year-olds	1039 (20)	1126 (23)	1093 (13)
5-year-olds	1313 (26)	1477 (27)	1417 (22)
Item-duration	556 (16)	511 (15)	487 (15)

This pattern was repeated for the 8-year-olds. There was a significant effect of Word-Type for decision latencies (F2[2,68]=12.36, p<0.001). Again, pairwise comparisons showed significant differences between Word and Assimilated items (F2[1,44]=17.16, p<0.001) and between Word and Nonword items (F2[1,46]=10.37, p=0.002), but not between Assimilated and Nonword conditions (F2[1,45]=0.20). As for the 5-year-olds there were no differences in errors between conditions (F2[2,68]=1.31).

Table 2: Percentage Error Rates for Experiment 1. (standard errors are given in parentheses).

	Word	Assimilated	Nonword
Adults	5.7 (1.9)	2.5 (1.1)	2.6 (0.8)
8-year-olds	7.4 (1.1)	3.3 (1.1)	5.5 (1.9)
5-year-olds	9.9 (2.1)	10.6 (2.4)	9.0 (1.7)

The adults showed the same pattern of results as the children. There was a main effect of Word-Type for decision latencies (F2[2,69]=6.65, p=0.002). Pairwise comparsions showed significant differences between Word and Assimilated items (F2[1,45]=7.85, p=0.007) and Word and Nonword items (F2[1,47]=10.93, p=0.002), but not between the Assimilated and Nonword items (F2[1,45]<1). The adults showed a marginal difference in error rates between conditions by-subjects (F1[2,24]=3.06,p=0.07), but not by items (F2[2,68]=2.14). This was due to the error rate in the Word condition.

Both the 5-year-old and 8-year-old children and the adults rejected the assimilated forms and the nonwords equally easily. This suggested that by 5-years-old children have a level of representational detail in their URs which is in line with the Full-Specification hypothesis. This leads to an adult-like level of perceptual intolerance to mismatch between surface forms and stored lexical representations. These results also suggest that adult's URs are fully-specified which seems to be inconsistent with the findings of Lahiri and Marslen-Wilson (1991). But there is other evidence that adults do not accept assimilated SFs in isolation as tokens of URs (Marslen-Wilson, Nix and Gaskell, 1995).

Experiment 2

Experiment 1 suggested that even 5-year-old's URs are fully-specified. However, this experiment only looked at the processing of single words. This pattern of behaviour might be different in continuous speech, not least because assimilated forms result from processes occurring in connected speech. Following Gaskell and Marslen-Wilson (1996) we investigated children's sensitivity to the viability of assimilation as a means of probing word recognition in continuous speech. Two groups of children[2], 32 6-year-olds (mean age = 6;02, range: 5;07 to 6;09) and 33 9-year-olds (mean age = 9;06; range: = 8;09 to 9;09), and a group of 51 adults participated in a word monitoring task.

Participants heard sentences and were asked to listen out for a target word and make a button-press response when it occurred. The test word immediately preceding the target was either a canonical form, a viable place assimilation, an unviable place assimilation or a nonword (a single feature distant from a word). If the test word disrupted processing in any way, this would show up in slowed monitoring latencies relative to the baseline condition, in this case the unchanged word.

Forty pairs of sentences were constructed, consisting of a context sentence and a test sentence[3] as follows (the capitalised word is the target for monitoring): Unchanged: "Mary put her hand up for the teacher. The *hard* BOOK was too difficult for her to read."; Viable assimilation: "...The [hɑ:b] BOOK was..."; Unviable assimilation: "...The [hɑ:g]BOOK was..."; Nonword: "...The [hɑ:z] BOOK was...".

The targets were always pictureable nouns and the test words were either verbs or adjectives[4]. The syntactic predictability of the targets was reduced by the use of adjectives and verbs as test words, and by placing the targets in different clauses in the test sentence. To ensure nonwords were not cues to targets, filler sentences were included in which nonwords were not adjacent to targets. The sentences were rotated into four versions of the experiment so that each form of the test sentences only occurred once in each version.

Again the contrasting views of lexical representation make different predictions. Gradual Segmentation predicts that the younger children will not be sensitive to the single feature deviations in the test words and so their word monitoring should not be disrupted in any of the changed conditions relative to the Unchanged condition. Radical Underspecification predicts word monitoring should not be

---

[2] Although the children in Experiment 2 were slightly older than those in Experiment 1, they were drawn from the same academic year group.

[3] The sentences were pretested on adults participants using a sentence completion task to ensure that none of the target words were predictable from the preceding context.

[4] These words were chosen from words within the productive vocabularies of 2- to 5-year-olds from CHILDES corpora, as in Experiment 1.

disrupted in either of assimilated conditions, because the changed place features are not specified lexically. They will, however, be disrupted by the Nonword condition, because of a mismatch on a lexically specified feature. Full-Specification predicts disruption in the Nonword condition because of a mismatch on the lexically specified changed feature. Performance on the assimilated conditions will depend on the ability to do phonological inference. Canonical URs could be consistent with perceptual intolerance and the acceptability of variant SFs, if a phonological parser is assumed to operate in continuous speech processing which can apply phonological rules to SFs to derive URs before the lexicon is accessed (Church, 1987). If children can do phonological inference, or parsing, they will be disrupted in the Unviable condition, but not the Viable condition. If they cannot, they will be disrupted in the both the assimilated conditions because they will be unable to derive the appropriate UR.

Results and Discussion

The mean monitoring latencies for Experiment 2 are given in Table 3.

Table 3: Monitoring Latencies for Experiment 2 in msec (standard errors are given in parentheses).

	Unchanged	Viable	Unviable	Nonword
Adults	316 (8)	314 (8)	334 (9)	375 (9)
9-year-olds	342 (14)	351 (10)	366 (12)	394 (11)
6-year-olds	461 (17)	466 (15)	456 (15)	494 (12)

Adults performed in a way that was consistent with the results obtained by Gaskell and Marslen-Wilson (1996). There was an overall significant effect of Word-Type ($F1[3,129]=39.86$, $p<0.001$; $F2[3,117]=16.59$, $p<0.001$). Bonferroni corrected pairwise comparisons showed there was a significant difference between the Unchanged and Unviable conditions ($t1[46]=2.92$, $p=0.003$; $t2[39]=1.83$, $p=0.038$), the Unchanged and Nonword conditions ($t1[46]=7.64$, $p<0.001$; $t2[39]=6.44$, $p<0.001$), the Viable and Unviable conditions ($t1[46]=3.34$, $p=0.001$; $t2[39]=2.19$, $p=0.018$) and the Unviable and Nonword conditions ($t1[46]=5.66$, $p<0.001$; $t2[39]=3.92$, $p<0.001$). In contrast, no difference was found between the Unchanged and Viable conditions ($t1[46]<1$; $t2[39]<1$). This indicated that the canonical SFs and assimilated SFs in a licensing phonetic context, were equally acceptable, showing evidence of functional underspecification. We also found that encountering an unviable SF was less disruptive than encountering a genuine nonword. The explanation for this finding may lie in the fact that unviable SFs only mismatch with the lexically stored target representation after initial lexical access as a result of phonological inferencing (Gaskell and Marlsen-Wilson 1996), whereas the Nonwords will fail to access any lexical representation.

The pattern of results found with the 9-year olds was broadly similar to that of the adults, although there were some differences. There was an overall significant effect of Word-Type ($F1[3,78]=13.54$, $p<0.000$; $F2[3,114]=6.74$, $p<0.001$). The pairwise comparisons for 9-year-olds showed significant differences between the Unchanged and Nonword conditions ($t1[29]=5.05$, $p<0.000$; $t2[38]=3.98$, $p<0.000$), the Unchanged and Unviable conditions ($t1[29]=2.85$, $p<0.004$; $t2[38]=2.15$, $p<0.019$), and the Unviable and Nonword conditions ($t1[29]=2.98$, $p=0.003$; $t2[38]=2.08$, $p<0.044$). The difference between the Unchanged and Viable conditions was not significant ($t1[29]=1.68$, $p<0.052$; $t2[38]=1.39$, $p<0.086$). However the difference between the Viable and Unviable conditions was at best a nonsignificant trend ($t1[29]=1.90$, $p=0.034$; $t2[38]=1.10$, $p=0.140$).

The 9-year-olds were disrupted by the Nonwords, which does not necessarily contradict Gradual Segmentation as they should have developed segmented representations by this age. These results are inconsistent with the notion that children are using fully-specified URs and a phonological parser to translate between SFs and URs, because the lack of a clear difference between the Viable and Unviable conditions suggests that phonological parsing is not operating, in which case both SFs should mismatch with the fully-specified UR; but the Viable and Unchanged conditions did not differ. A possible explanation for this pattern of results is that the children's URs are underspecified. The pattern might be consistent with underspecified URs and no phonological inference, but the difference between the Unchanged and Unviable conditions is evidence of phonological inference operating. The overall pattern of results suggested that the 9-year-olds were converging on adult-like performance.

The results for the 6-year-olds helped to clarify the developmental picture. Whilst there was an overall main effect of Word-Type ($F1[3,78]=3.61$, $p=0.017$; $F2[3,117]=2.61$, $p<0.055$), the pairwise comparisons only showed up significant differences between the Unchanged and Nonword conditions ($t1[29]=2.69$, $p=0.006$; $t2[39]=2.31$, $p=0.013$) and the Unviable and Nonword conditions ($t1[29]=2.16$, $p=0.039$; $t2[39]=2.30$, $p=0.027$). The other comparisons, between Unchanged and Viable conditions ($t1[29]<1$; $t2[39]<1$), Unchanged and Unviable conditions ($t1[29]<1$; $t2[39]<1$) and Viable and Unviable conditions ($t1[29]<1$; $t2[39]<1$), were not significant.

The 6-year-olds were disrupted by the Nonwords which suggested their URs were not more holistic than those of older children or adults. However, they were insensitive to the phonological viability of assimilations which would be consistent with Radical Underspecification, on the assumption 6-year-olds achieve lexical access because underspecified representations are the perceptual targets of the process. But this explanation is contradicted by Experiment 1, which showed that 5-year-olds would not accept assimilated forms in isolation as real words. These results are more consistent with the children's lexical representations being functionally underspecified and with phonological inference having as yet to develop. Indeed, the evidence from the 9-year-olds suggested that it is only beginning to emerge at that age.

Computational Modelling

The experimental evidence presented above suggested a dissociation between the development of abstract lexical representations and phonological inference. In the approach taken by Gaskell et al. (1995) both functional underspecification and phonological inference emerge as a result of the mapping between variant SFs and the lexicon. At first sight this linkage would not seem to provide an explanation of the dissociation found here. In order to look into the possibility that phonological inference develops slowly relative to lexical development, we trained a series of connectionist networks to learn these processes. The simulations were run using an artificial language which allowed the modelling to focus on the effects of phonological variation to the exclusion of other factors. The networks were trained on two parallel tasks. The first was learning a small vocabulary, by mapping between a featural representation of phonemic segments and localist lexical representations. The second was to mark viable assimilations by activating a single output unit concurrently with the initial segment of the following word (i.e. the context for the assimilation). This unit represented the model's learning of phonological inference with a mapping which was comparable to the lexical one. Its psychological analogue can be seen as the information available during lexical access when the match between SFs and URs is evaluated in the light of following context.

The simple recurrent networks used consisted of an 8-unit input layer, one corresponding to each of the 8 features used to code phonemes, feeding forward to 50 hidden units with recurrent connections to 50 context units, allowing the network to internally represent the sequential dependencies in its training environment. The hidden units were connected to an output layer which consisted of 30 localist lexical units and a single phonological inference unit.

The artificial language used had a phonology of 13 consonants and 4 vowels with each phoneme represented as a bundle of 8 phonetic features. The phonology of the language allowed place assimilation to occur. A vocabulary of 30 CVC words was constructed from this phonology. Ten of the words ended in coronal stops, and so were assimilable, and another 20 began with noncoronal stops, providing a context for assimilation. Localist representations were used to distinguish between lexical items.

The networks were trained using standard backpropagation. An epoch of training consisted of a single sweep through 100 occurrences of each item presented in random order, and the training set was re-randomised after every epoch. The networks were trained at three rates of assimilation: 100% assimilation, where assimilation occurred in all viable contexts, 50% assimilation and 25% assimilation. This in effect progressively reduced the token frequency of assimilated segments from 22% to 11% and then to 6%, respectively. At 100% assimilation, assimilated SFs were more common than canonical the forms of assimilable words. Three networks were run at each rate of assimilation.

Whilst our primary focus was to compare the rate of learning for the assimilated forms with the learning of phonological inference, we were also interested in the way the network would deal with random variation, and so a set of nonwords was created by taking the assimilable items and changing a single feature of the word-final segment in a phonologically random way. The measure of the network's performance was simply the activation level of the output nodes, which in this localist model gave a direct measure of the goodness-of-fit between input and the model's internal representations. The networks were tested on randomised sets of 4 tokens of each of the words in its vocabulary and the nonwords.

Results and Discussion

The results for each rate of assimilation were averaged over three networks. The learning curves for assimilated words and phonological inference and the model's response to nonwords at different rates of assimilation are presented in Figures 1 to 3. Output activations were measured at word-offset for lexical units and word-onset for the inference unit. Figures 1-3 show mean output activations for each set of test items at different points in the model's development.

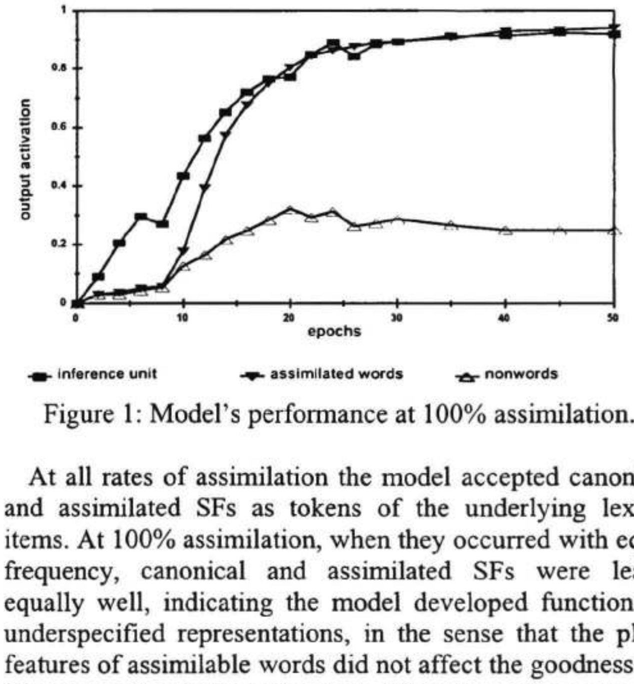

Figure 1: Model's performance at 100% assimilation.

At all rates of assimilation the model accepted canonical and assimilated SFs as tokens of the underlying lexical items. At 100% assimilation, when they occurred with equal frequency, canonical and assimilated SFs were learnt equally well, indicating the model developed functionally underspecified representations, in the sense that the place features of assimilable words did not affect the goodness-of-fit computation in lexical access. The pattern for nonwords was similar for all rates of assimilation. Only very early in training, up to around 10 epochs, did nonwords achieve similar levels of activation as the assimilated forms, indicating the model acquired an intolerance to random variation early on in training.

The model also learnt to do phonological inference, as indicated by inference node activation in the presence of viable assimilations. At 100% assimilation, the inference task was learnt more easily than the lexical mapping for assimilated forms, but at the lower rates of assimilation the rate of learning for the inference task dropped off markedly. In particular, the relative delay between the model's ability to recognise viable assimilations and its ability to recognise assimilated forms increased sharply as the rate of assimilation dropped from 50% to 25%. Thus, at low token frequencies for assimilated SFs a clear dissociation emerged between the model's development of abstract representations and phonological inference.

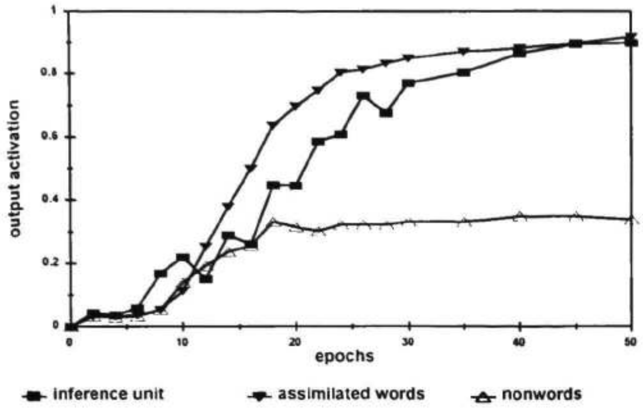

Figure 2: Model's performance at 50% assimilation

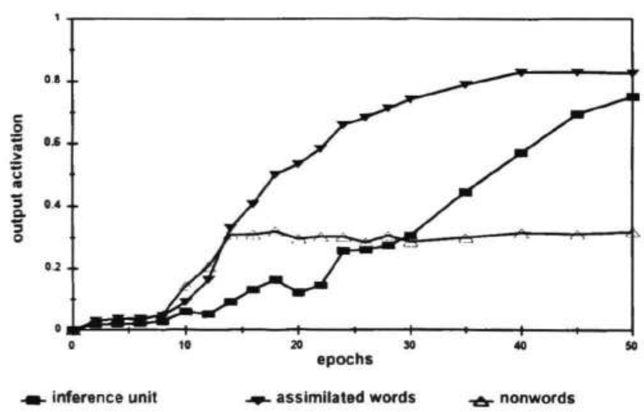

Figure 3: Model's performance at 25% assimilation

Conclusions

We identified three theories of lexical representation in children, Full-Specification, Radical Underspecification and Gradual Segmentation. The results of Experiments 1 and 2 suggested that none of these approaches captured the pattern of performance seen in children between the ages of 5- and 9-years-old. From at least 5-years-old, children showed evidence that their lexical representations were specified to the level of single phonetic features but that they were also functionally underspecified because assimilated SFs did not mismatch with URs in continuous speech processing. Whilst the perceptual targets for lexical access available to children are representationally adult-like, their processing differs. Six-year-olds accepted assimilated forms regardless of the viability of the assimilation, and 9-year-olds' performance suggested an emerging sensitivity to phonetic context. A computational model of the development of spoken word recognition suggested the reason for the dissociation between the development of functionally abstract representations and phonological inference results from the different effects of the frequency of assimilated forms on these two processes. Lexical learning for assimilated forms can proceed with every exposure to a word, regardless of its surface form. In contrast, learning the relationship underlying place assimilation is entirely dependent on occurrences of assimilation. When the relevant information is sparse, learning this structural relationship is attenuated.

Acknowledgments

This research was supported by MRC research studentship RF-95-102 and an MRC Programme Grant to W.D. Marslen-Wilson and L.K. Tyler.

References

Archangeli, D. (1988). Aspects of underspecfication theory, *Phonology*, 5, 183-207.

Church, K.W. (1987). Phonological parsing and lexical retrieval. *Cognition*, 25, 53-69.

Elman, J. (1990). Finding structure in time. *Cognitive Science*, 14, 179-221.

Gaskell, M.G. & Marslen-Wilson, W.D. (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Memory and Performance*, 22(1), 144-158.

Gaskell, G.M., Hare, M. & Marslen-Wilson, W.D. (1995). A connectionist model of phonological representation in speech perception. *Cognitive Science*, 19, 407-439.

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

Gierut, J.A. (1996). Categorisation and feature specification in phonological acquisition. *Journal of Child Language*, 23, 397-415.

Lahiri, A. & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition*, 38, 245-294.

Marslen-Wilson, W. (1993). Issues of process and representation in lexical access. In G.T.M. Altmann & R. Shillcock (Eds*)*, *Cognitive Models of Speech Processing: The Second Sperlonga Meeting*. Hove: Lawrence Earlbaum Associates.

Marslen-Wilson, W., Nix, A., & Gaskell, G. (1995). Phonological variation in lexical access: abstractness, inference and English place assimilation. *Language and Cognitive Processes*, 10, 285-308.

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

Werker, J.F., & Lalonde, C.E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology*, 24, 672-683.