In Mozer, M. C., Smolensky, P., Touretzky, D. S., Elman, J. L., and Weigend, A. S. (Eds.). (1994). *Proceedings of the 1993 Connectionist Models Summer School*. Hillsdale, NJ: Lawrence Erlbaum.

Investigating Phonological Representations: A Modeling Agenda

Prahlad Gupta
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15217
prahlad@cs.cmu.edu

This paper outlines a bottom-up research approach to studying lexical representation, emphasizing the development of infant phonological perception as a source of data that can illuminate the nature of phonological representations. It is proposed that neural network formalisms can provide a useful framework for thinking about these issues. We identify a key set of empirical results, understanding of which would yield considerable insight; this defines a modeling agenda for investigation of phonological representations. We report preliminary simulations that explore this agenda, exemplifying how network modeling techniques can contribute to understanding of these phenomena.

MOTIVATION

What is the nature of words? How are they represented mentally? What are the properties of these representations? How are these representations evoked in processing speech?

These issues clearly have importance for thinking about language. For example, conceptions of lexical access in sentence processing will differ considerably depending on whether lexical representations are conceived of as being "static" or "dynamic", or whether they are viewed as being "distributed" or "localist". Investigating the mental representation of lexical form can therefore be viewed as part of a "bottom-up" strategy to studying language. Such investigation is also bottom-up in the sense that the development of representations of words necessarily must precede the development of more complex language abilities in the human child.

It seems uncontroversial to suppose that the formation of representations is a major part of what goes on in the early stages of language acquisition, especially over the first two years of life. As Aslin [1] has noted, it seems unlikely that infants actually hear speech as meaningful units until sometime in later infancy when they begin to associate sounds with meanings. It is therefore reasonable to think of these earliest representations as denoting the infant's perception of speech sounds, relatively independent of their meaning.

Consequently, it seems valuable to enquire into the nature of such *phonological* representations, and into the processes by which they might develop: these enquiries would reach into the very heart of early language development. What, then, is the nature of phonological representations? How do they develop? And how are they involved in the processing of spoken words? Important as these issues are, little is known about them, as Lahiri & Marslen-Wilson [13] have pointed out.

The hypothesis advanced in this paper is that neural network formalisms can provide a framework for thinking about phonological perception and its development, and thus can contribute to understanding the nature of phonological representation and processing[1]. The first section of the paper outlines a specific set of developmental phonological perception phenomena, proposing these as a test-bed for the above hypothesis. The second section of the paper reports on some preliminary neural net simulations that explore a subset of these phenomena; the results support the hypothesis that network formalisms may be an illuminating framework for thinking about these issues.

TOUCHSTONE PHENOMENA FOR PHONOLOGICAL PERCEPTION

This section presents and discusses selected empirical results in infant phonological perception, obtained over approximately the first one year of life. It will be argued that these results provide a testbed for the hypothesis that network formalisms can provide a framework for thinking about phonological perception and its development, and thus constitute an agenda for modeling.

Categorical perception. The discrimination functions of human adults listening to phonetic segments varying along an acoustic continuum (such as voice-onset time, or place of articulation) show peaks of discriminability corresponding to the locations of phonemic category boundaries as determined by absolute identification experiments [11]. This phenomenon, viz., the occurrence of steep crossovers for both identification and discrimination, has been termed *categorical perception*. It occurs only for consonants; vowels appear to be perceived continuously [1]. In adults, categorical perception occurs only for sounds in the adult's native language [11]. In human infants, categorical discrimination has been demonstrated as early as 1 month of age: the infants reacted as if they perceived a sudden shift in the series of phonetic (consonantal) segments along an acoustic continuum, at the same point as the adult-defined boundary between two phonemic categories. These effects have been obtained along various acoustic continuua, and it appears that infants can discriminate virtually every speech contrast used in one of the world's languages; it also seeems likely that prior to 6 months of age, infants are performing their analysis of speech sounds solely on the basis of acoustic differences, which are sufficient to permit categorical discrimination [1].

The perceptual magnet effect. Kuhl's work [10] indicates that adult listeners' ratings of the goodness of exemplars of a vowel vary, even while all exemplars are categorized as being the same vowel. Different exemplars of /i/ were rated by adults, and received different ratings of "goodness" (which were very consistent across raters), although always being categorized as /i/. This suggests that there are "prototypically good" vowels.

Adults also show an asymmetry in *discrimination* of "prototypical" vs. "nonprototypical" stimuli. When adults are presented with the "best" exemplar of a vowel repeatedly and then a "peripheral" exemplar of that same vowel, they often fail to discriminate the difference. However, when presented with peripheral exemplars and then tested with the "best" exemplar, discrimination is better. This is what Kuhl [10] has called the *perceptual magnet effect*. According to her, the central member of the category seems to "capture" the other instances, rendering them less discriminable.

Kuhl has also shown that infants' responses to protoypical vs. nonprototypical vowel stimuli correspond with adults' goodness ratings. Stimuli were synthesized to form four concentric rings around each of two central stimuli: an adult-rated prototypical, and an adult-rated non-protoypical /i/. Infants heard one or other of the central stimuli as the reference stimulus, and were tested on discrimination of the surrounding stimuli. Results from infants responding to these stimuli were as follows:

(1) Stimulus generalization for both groups, i.e., generalization of the head-turn response from the center stimulus to the surrounding stimuli. (2) A group effect, with generalization at a given distance from the central stimulus

---

[1] There are, of course, existing accounts of phonological perceptual development [8, 21], and the present approach draws importantly on many of their ideas. However, these previous approaches have not focussed on computational instantiation of the frameworks they propose. Other recent work does examine phonological development from a computational standpoint [14], but is not primarily concerned with the nature of representations or with accounting for specific perceptual phenomena.

being higher for the prototype-based group than for the non-prototype-based group. This is consistent with the "perceptual magnet" effect, since greater generalization from the prototype is the same thing as poorer discrimination from the prototype (demonstrated with adults). (3) One vector of stimuli was actually shared between the two concentric rings of stimuli. The two groups of infants were therefore both tested on this set of stimuli, but in opposite "directions". Infants exposed to the prototype as the reference stimulus discriminated only the most distant stimulus along this vector, whereas infants exposed to the non-prototype as reference discriminated stimuli from the second nearest on. This is a direct replication of the adult perceptual magnet effect. (4) There was a .95 correlation of adults' goodness ratings and infants' generalization scores around the prototype.

Language-specificity of the perceptual magnet effect. In further work by Kuhl [12], one set of vowels was synthesized around a prototypic exemplar of English /i/ and another set was synthesized around a prototypic exemplar of Swedish /y/. For English adults /y/ is not perceived as a prototype of any English vowel, and for Swedish adults /i/ is not perceived as prototypical of any Swedish vowel.

6-month old English- and Swedish-learning infants were tested for discrimination from both the central stimuli. English infants showed a stronger generalization from the English /i/ stimulus to its variant stimuli, than from the /y/ stimulus to its surrounding stimuli. The converse was true for the Swedish infants. This suggests that the prototype structure for vowels has begun to be tuned by the ambient linguistic environment at least by 6 months of age.

Discriminability of nonnative contrasts and its development: consonants. Cross-language studies of consonant discrimination have revealed that, up to about 6 months of age, infants can discriminate nearly every phonetic contrast on which they have been tested, including contrasts comprised of phonetic segments that are not phonemically contrastive in their language-learning environment (but that are contrastive in some other human language).

Werker [21] and others have shown, however, that between 8 and 12 months of age, infants seem to lose the ability to discriminate between the same nonnative contrasts that they could earlier discriminate. This loss of nonnative contrasts only appears to take place for phonetic segments that are assimilable to a native language category. Infants retain the ability, however, to discriminate native language contrasts.

Adult speakers are unable to discriminate many nonnative contrasts that infants can discriminate at earlier ages, but lose by 12 months. They are able to discriminate the contrasts that infants continue to be able to discriminate. This is consistent with the notion that nonnative segments are being assimilated to a native language category (if a sufficiently close one exists), in both the older infants, and the adults. Adult discrimination performance can, however, improve with training or practice.

Discriminability of nonnative contrasts and its development: vowels. In further work by Werker and colleagues [16], English-learning infants were tested on their ability to discriminate two pairs of high front-rounded vs. high back-rounded German vowels. All vowels were presented in the context [dVt]. One contrast involved tense high front vs. high back (/u/ vs. /y/) and the other involved lax high front vs. high back (/U/ vs. /Y/). Both contrasts are phonemic in German, but not in English.

Adult English speakers were as good as German speakers at discriminating between members of both contrastive German vowel pairs. However, each back German vowel was perceived to be more like an English vowel than was the front German vowel; in this sense, each German back vowel was designated by the experimenters as the more "prototypical" member (for English adults) of each contrasting German pair.

4-month-old English-learning infants were able to discriminate between members of both contrastive German vowel pairs, irrespective of whether the more or less prototypical member served as the background category. 6- to 8-month-old English-learning infants were able to discriminate the German contrasts when the background stimulus was the non-prototypical member, but not when it was the prototypical member. By 10-12 months, English infants no longer discriminated the German contrasts, irrespective of the "direction" of testing.

The above results can be summarized as follows. (1) First, the perceptual magnet effect seems to exist for vowels, but not for consonants. (2) There is also another basic difference in adult perception of vowels vs. consonants. For phonemically contrastive consonant sounds in the listener's native language, there are steep crossovers for both labeling and discrimination. For nonnative listeners, category labeling boundaries are absent, less sharp, or do not coincide with those of native listeners. Adults have considerable difficulty discriminating certain nonnative constrasts. The situation is slightly different for vowels: as with consonants, there are steep crossovers for labeling of phonemically contrastive vowels in the listener's native language. Discrimination, however, is more continuous, for both native and nonnative listeners, and in fact the discrimination abilities of nonnative listeners are typically quite good compared with those of native listeners. However, these discrimination abilities appear subject to the "perceptual magnet" effect. (3) Infants are initially sensitive to virtually any native or nonnative phonemic contrast, whether between vowels or consonants, up to about 4 months of age. (4) Beyond this age, their sensitivities seem to get attuned to environmental (native) language sounds, leading to loss of sensitivity to nonnative contrasts by 6-8 months for vowels, and by 10-12 months for consonants. (5) For vowels, there seems to be an intermediate point (around 6-8 months) in this loss of sensitivity. At this intermediate point, discriminatory ability is preserved only if the nonnative sound serving as reference stimulus is *not* prototypical, i.e., similar to a native language sound.

These data represent some of the central findings in infant speech perception over the last four decades, and constitute a key and inter-related set of phenomena. An important aspect of the phenomena is their developmental nature, as well as the developmental progression in which they occur. Providing a computational account of these data would yield considerable insight into fundamental questions regarding the nature of phonological representation. This set of phenomena is therefore proposed here as a testbed for investigation of the nature and development of phonological representations via network modeling techniques, in much the same way that a miniature language acquisition problem has been proposed by Feldman and colleagues [2] as a touchstone for cognitive science.

The next section describes simulations that begin to explore this modeling agenda; the results demonstrate the role that neural network techniques can play in thinking about phonological representation and processing.

SIMULATIONS: LOSS OF NONNATIVE CONTRASTS

As discussed above, infants lose sensitivity to nonnative contrasts towards the end of the first year of life. In particular, it has been found by Werker and colleagues [21] that English-learning infants aged 6-8 months are able to discriminate Hindi and certain other nonnative contrasts, while infants aged 10-12 months are mostly unable to do so, as are adult native speakers of English. However, English-learning infants at all ages, as well as adults, retain the ability to discriminate certain other nonnative contrasts, such as that between two Zulu clicks.

Part of Werker's account of these phenomena is that both sounds in the Hindi contrast (involving a dental vs. a retroflex [ta]) may by the later age have become assimilated to the native English alveolar /ta/, and thus ceased to be discriminable. The Zulu clicks, on the other hand, may not be easily assimilable to any known category, and hence remain discriminable. The intuition is that the infant's mental perceptual landscape initially has a topology that allows for discrimination between virtually any speech stimuli; however, this topology is altered by exposure to the native language environment in such a way that nonnative distinctions become blurred and no longer discriminable.

In neural network terms, such properties might be expected to follow naturally from the development of *attractor states*. The idea here is that the energy landscape in the network initially has low-energy basins (i.e., attractors) corresponding to essentially each possible phonetic segment. However, learning through exposure to the native language re-sculpts this energy surface, and the attractor states that continue to exist are those that correspond to the phonemes of the native language.

To examine the ability of neural networks to flesh out these intuitions, we constructed simulations in which

b	1.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
a	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	1.0
i	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0
ba	1.0	1.0	1.0	0.0	0.0	0.0	0.0	0.5	1.0
bi	1.0	1.0	1.0	0.0	0.0	0.0	0.0	1.0	0.0

Figure 1: Input representation for phones and demisyllables.

networks' responses to "nonnative" stimuli were examined both before and after training on a set of "native" stimuli.

Input Representation and Data

The actual input set comprised a set of demisyllables, representing the "native language" sounds to which an infant might be exposed. Of course, a human infant is in reality not exposed to streams of syllables, but rather, to actual linguistic utterances. Gupta & Mozer [3] have shown how demisyllables can be extracted from words, using a simple mechanism based on autopredictive error and stress level, while Gupta & Touretzky [6, 4, 5] have examined the ability of a perceptron to assign stress to syllables. The details of this processing are not important for present purposes; we simply assume some such prior processing that extracts demisyllables from actual words, as a result of which the infant's phonological perceptions are structured in terms of demisyllables. There appears to be considerable agreement that this is the case with human infants [8]. Thus the input corpus we used comprised the syllables *ba be bi d@ di do g@ ga ge go k@ ka ki ko la ma mi mo pa pe pi po t@ ta ti to*.

To represent these syllables, we used a phonetic feature representation scheme proposed by Shillcock et al. [20], in which each possible phone is encoded in terms of a set of 9 feature values, which are intended to have physical correlates in the speech signal. These features are: (1) oral cavity openness, (2) palatality, (3) labiality, (4) occlusion, (5) aperiodic energy, (6) nasality, (7) apicality/coronality, (8) velarity/centrality, and (9) voicelessness [20]. Figure 1 shows the encoding of the phones [b], [a], and [i] in terms of these 9 features. Each bit of the 9-element vector shown in the figure represents the value for that phone on one feature dimension. To encode a syllable such as *ba*, we simply superimposed the 9-element vectors representing the two segments *b* and *a*, as also shown in Figure 1.

To examine the model's responses to unknown sounds, a "dental" and "retroflex" *t* were simulated by modifying the value of the "coronality" feature, from 1.0 for the alveolar *t*, to 0.7 and 0.3 for the dental and retroflex versions respectively. These were used to create "nonnative" stimuli: a dental *ta*, which will be denoted by *t(d)a*, and a retroflex *ta*, denoted by *t(r)a*. The syllables *na* and *nga* were treated as a second nonnative contrast, since neither of them was included in the input corpus.

Network Architecture and Processing

Three alternative architectures were investigated (Figure 2). These were (1) a Deterministic Boltzmann Machine (DBM) [15, 7], (2) a Competitive Learning network (CL) [19], and (3) a Multi-Layer Perceptron (MLP) [17]. In each case, the input to the network was the nine-element vector representing a demisyllable. For the DBM and MLP, the task of the network was to reproduce the input layer vector at the output layer (see Figure 2). For the CL architecture, the task was to categorize the input layer vector, by turning on exactly one of the output layer units.

For the DBM, the output activation is obtained by applying the input vector, and then performing synchronous updates of unit activations in repeated cycles until the magnitude of changes in unit activations falls below a specified criterion, i.e., until the network settles. The output unit activations at this time constitute the network's response. For the MLP, the output layer activation is produced by propagating the input vector forward in one pass. For the CL architecture, the "winner" is chosen to be the unit with weights closest to those of the input vector, as in the standard

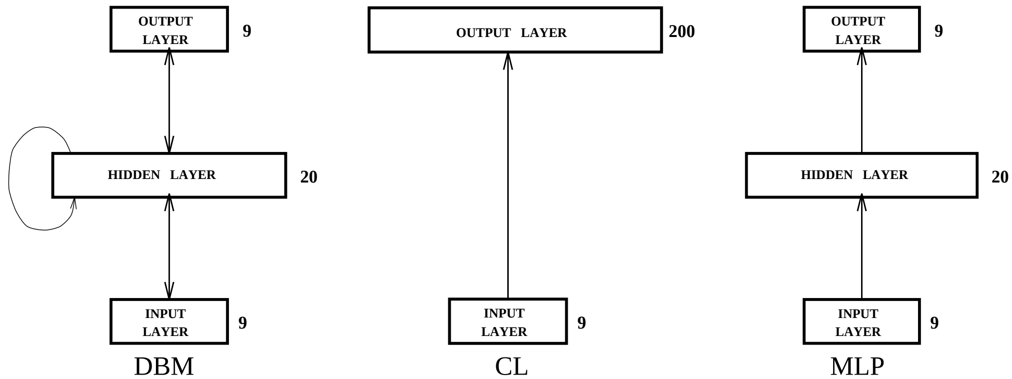

Figure 2: Classification network architectures. Deterministic Boltzmann Machine (DBM); Competitive Learning Network (CL); Multi-Layer Perceptron (MLP). Numbers indicate number of units in a layer. Arrows indicate connectivity.

Table 1: Responses of CL classification network to nonnative stimuli.

Before training				After training			
ta	153	*ma*	92	*ta*	153	*ma*	92
t(d)a	153	*na*	149	*t(d)a*	153	*na*	100
t(r)a	50	*nga*	204	*t(r)a*	153	*nga*	39

Note: The numbers identify the unit responding to a particular stimulus.

algorithm [9], but with the additional requirement that the error for the winner be below a specified criterion; if it is not, an "uncommitted" unit is chosen to be the winner. During training, this error criterion was progressively relaxed.

Weight adjustment for the autopredictive network and MLP classification network was via the back-propagation algorithm [17]. Weights in the DBM classification network were adjusted via contrastive Hebbian learning [7]. In the CL classification network, the winner's weights were adjusted via the standard competitive learning equation [9].

Simulation of Loss of Sensitivity to Nonnative Contrasts

The networks were tested on the nonnative stimuli prior to any training on the input corpus. They were then trained on the input corpus for 20 to 100 epochs, and then tested again on the nonnative stimuli.

For nonnative stimuli *before* training, results from the CL classification network are shown in the left-hand part of Table 1. The numbers are merely identifying labels for the unit responding to a particular stimulus. As shown, the nonnative stimuli are responded to by different units, indicating their discriminability [2]. Results from the CL classification network *after* training are shown in the right-hand part of Table 1. The same unit now responds to *ta, t(d)a,* and *t(r)a*, indicating that the nonnative stimuli have been assimilated to the known *ta* category. The *na* and *nga* stimuli are still responded to by different units, however, indicating that they are not assimilable to known categories.

Equivalent results were obtained with the DBM, but not the MLP architecture. Output responses of the DBM classification network are shown projected onto the first two principal components, before training (Figure 3a). The

---

[2] Both *ta* and *t(d)a* are responded to by the same unit, suggesting that these stimuli are already "perceived" as similar. More importantly, however, the two nonnative stimuli *t(d)a* and *t(r)a* are perceived as distinct.

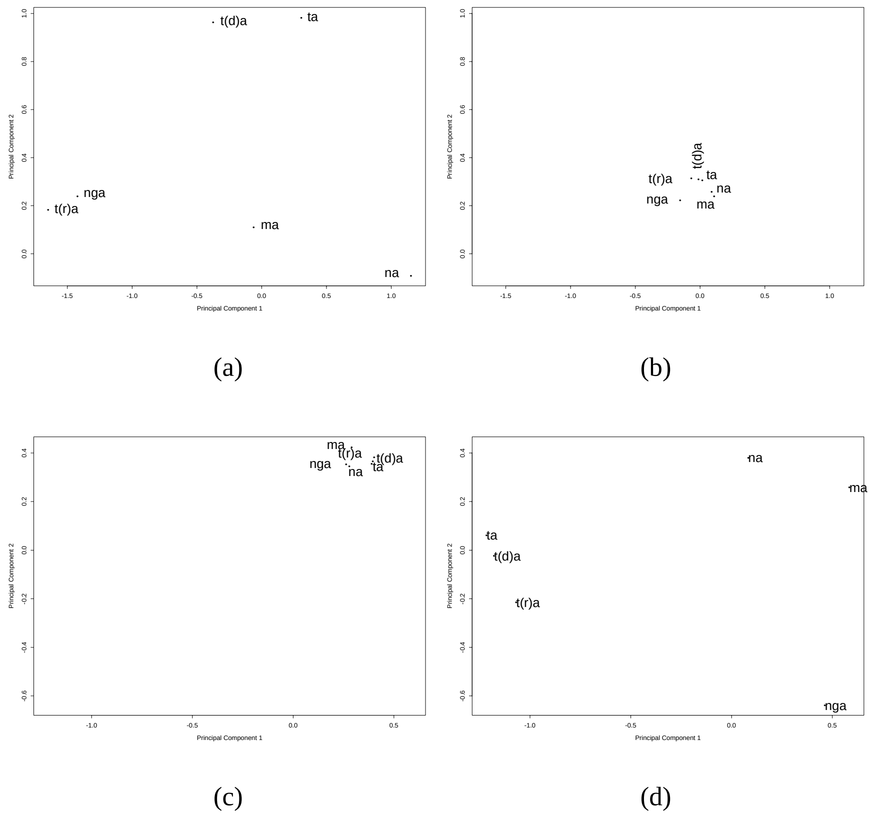

Figure 3: Responses of DBM and MLP classification networks to nonnative stimuli before and after training, projected onto the first two principal components. (a) DBM before training. (b) DBM after training. (c) MLP before training. (d) MLP after training.

network's responses are quite widely separated in state space, indicating discriminability of all the stimuli. After training, however (Figure 3b), responses to the stimuli are much less dispersed in state space. Note, however, that *na* and *nga* are considerably further dispersed than are *t(d)a* and *t(r)a*. These results are analogous to those obtained with the CL architecture. With the MLP architecture, however, the opposite trend appears: responses to the stimuli are more widely dispersed after training (Figure 3d) than before training (Figure 3c).

These results can also be examined in terms of the average pairwise distance between members of the *ta–t(d)a–t(r)a* and *ma–na–nga* triples. With the DBM, the ratio of this average distance *after* training to the average distance *before* training was 0.42 for the stops, and 0.57 for the nasals, illustrating that discriminability had decreased for both groups, but more so for the stops. With the MLP architecture, however, the after-before ratio was 2.77 for the stops and 15.01 for the nasals, indicating that the members of each group had become *more* discriminable after training.

The results obtained with the CL and DBM architectures demonstrate lost sensitivity to certain nonnative contrasts as well as retained sensitivity to certain other nonnative contrasts. This models the observed developmental phenomena, and also provides a computational account of such a process, and thereby a basis for understanding why the observed selective loss of nonnative contrasts in infants might arise. As the perceptual ("classification") system develops, it becomes attuned to, and begins to categorize, sounds occurring in the environment. Other (nonnative) sounds now tend to be interpreted in terms of the categories developed for known, occurring sounds.

We began by hypothesizing that the computational notion of *attractor states* might aid in understanding the phenomena of interest. Pursuing this idea, we now consider the properties of the various classification network architectures examined. First of all, attractor states *necessarily* develop in the DBM, by virtue of its network dynamics. Its learned states thus represent basins of attraction; and this means that inputs similar to those that have been learned will tend to result in one of these attractor states. Second, the CL classification network approximates this property of the DBM, in that an input is mapped to the output unit with most closely similar weights, that is, in virtue of the output rule by which a winner is selected. Third, in a purely feedforward MLP, something like attractor states can develop under certain training regimes. For example, if an MLP is trained to categorize its inputs by turning on particular output units, then the weights developed in this task, together with an output interpretation procedure, will yield input-output mappings that have some of the properties of a system with attractor dynamics, in that many inputs will map onto a particular output state, and in that a given input will be classified by the most activated output unit. However, the MLP in the present case was trained to reproduce its input; such a training regime would not be expected to induce attractor-like properties.

The fact that the loss of nonnative contrasts is simulated with the CL and DBM architectures, but not the MLP architecture is therefore interesting, suggesting that the formation of attractor states is necessary to simulate this developmental trend. In the present simulations, the MLP does not form attractors, and is therefore unable to capture this phenomenon. Although these results need further investigation, they provide preliminary support for the hypothesis that attractor states may be a valuable notion in understanding phonological perception.

This simulation thus yields an interesting new way of thinking about phonological representations: as attractor states that are sculpted in perceptual space by exposure to language-specific input. This not only provides computational specification to the intuition we began with, it also provides a demonstration of how application of network modeling techniques to the set of "touchstone" phenomena proposed in this paper may begin to provide greater understanding of the nature and development of phonological representations.

Acknowledgements

I would like to thank Mike Mozer for collaboration on the simulation work reported in this paper, Janet Werker for bringing the perceptual magnet effect to my attention, Mike Mozer and Dave Plaut for detailed discussion of the ideas outlined here, and Jeff Elman, Brian MacWhinney, Jay McClelland, and Dave Touretzky for helpful comments at various points. Any inaccuracies or inconsistencies are, of course, solely my responsibility.

References

[1] R. N. Aslin, "Visual and auditory development in infancy," in *Handbook of Infant Development* (J. D. Osofsky, ed.), New York: Wiley, 1987.

[2] J. A. Feldman, G. Lakoff, A. Stolcke, and S. H. Weber, "Miniature language acquisition: A touchstone for cognitive science," Report TR-90-009, International Computaer Science Institute, Berkeley, CA, April 1990.

[3] P. Gupta and M. C. Mozer, "Exploring the nature and development of phonological representations," in *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society*, (Hillsdale, NJ), Lawrence Erlbaum, 1993.

[4] P. Gupta and D. S. Touretzky, "A connectionist learning approach to analyzing linguistic stress," in *Advances in Neural Information Processing Systems 4* (J. Moody, S. Hanson, and R. Lippmann, eds.), (San Mateo, CA), pp. 225–232, Morgan Kaufmann, 1992.

[5] P. Gupta and D. S. Touretzky, "Connectionist models and linguistic theory: Investigations of stress systems in language," *Cognitive Science*, in press.

[6] P. Gupta and D. S. Touretzky, “What a perceptron reveals about metrical phonology,” in *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*, (Hillsdale, NJ), pp. 334–339, Lawrence Erlbaum, 1991.

[7] G. E. Hinton, “Deterministic Boltzmann learning performs steepest gradient descent in weight-space,” *Neural Computation*, vol. 1, pp. 143–150, 1989.

[8] P. W. Jusczyk, “From general to language-specific capacities: The WRAPSA model of how speech perception develops,” Manuscript, 1992.

[9] T. Kohonen, *Self-Organization and Associative Memory*. Berlin: Springer-Verlag, 1984.

[10] P. K. Kuhl, “Human adults and human infants show a perceptual magnet effect for the prototypes of speech categories, monkeys do not,” *Perception and Psychophysics*, vol. 50, pp. 93–107, 1991.

[11] P. K. Kuhl, “Perception, cognition, and the ontogenetic and phylogenetic emergence of human speech,” in *Plasticity of Development* (S. E. Brauth, W. S. Hall, and R. J. Dooling, eds.), Cambridge, MA: MIT Press, 1991.

[12] P. K. Kuhl, K. A. Williams, F. Lacerda, K. N. Stevens, and B. Lindblom, “Linguistic experience alters phonetic perception in infants by 6 months of age,” *Science*, vol. 255, pp. 606–608, 1992.

[13] A. Lahiri and W. Marslen-Wilson, “The mental representation of lexical form: A phonological approach to the recogniion lexicon,” *Cognition*, vol. 38, pp. 245–294, 1991.

[14] L. Menn, K. Markey, M. Mozer, and C. Lewis, “Connectionist modeling and the microstructure of phonological development: A progress report,” in *Changes in Speech and Face Processing in Infancy: A Glimpse at Developmental Mechanisms of Cognition* (B. de Boyssons-Bardies, P. Jusczyk, P. MacNeilage, J. Morton, and S. de Schonen, eds.), Dordrecht, The Netherlands: Kluwer, in press.

[15] C. Peterson and J. R. Anderson, “A mean field theory learning algorithm for neural nets,” *Complex Systems*, vol. 1, pp. 995–1019, 1987.

[16] L. Polka and J. F. Werker, “Developmental changes in perception of non-native vowel contrasts,” Manuscript, 1993.

[17] D. Rumelhart, G. Hinton, and R. Williams, “Learning internal representations by error propagation,” in Rumelhart *et al.* [18].

[18] D. E. Rumelhart, J. L. McClelland, and the PDP Research Group, *Parallel Distributed Processing*, vol. 1: Foundations. Cambridge, MA: MIT Press, 1986.

[19] D. E. Rumelhart and D. Zipser, “Feature discovery by competitive learning,” in Rumelhart *et al.* [18].

[20] R. Shillcock, G. Lindsey, J. Levy, and N. Chater, “A phonologically motivated input representation for the modeling of auditory word perception in continuous speech,” in *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*, (Hillsdale, NJ), pp. 408–413, Lawrence Erlbaum, 1992.

[21] J. F. Werker and J. E. Pegg, “Infant speech perception and phonological acquisition,” in *Phonological Development: Models, Research and Implications* (C. Ferguson, L. Menn, and C. Stoel-Gammon, eds.), Parkton, MD: York Press, 1992.