

# Adults' self-directed learning of an artificial lexicon: The dynamics of neighborhood reorganization

Neil P. Bardhan (nbardhan@bcs.rochester.edu)

Richard N. Aslin (aslin@cvs.rochester.edu)

Michael K. Tanenhaus (mtan@bcs.rochester.edu)

Department of Brain & Cognitive Sciences, University of Rochester  
Meliora Hall, Box 270268  
Rochester, NY 14627 USA

## Abstract

Artificial lexicons have previously been used to examine the time course of the learning and recognition of spoken words, the role of segment type in word learning, and the integration of context during spoken word recognition. However, in all of these studies the experimenter determined the frequency and order of the words to be learned. Here we ask whether adult learners choose, either implicitly or explicitly, to listen to novel words in a particular order based on their acoustic similarity. We use a new paradigm for learning an artificial lexicon in which the learner, rather than the experimenter, determines the order and frequency of exposure to items. We analyze both the temporal clustering of subjects' sampling of lexical neighborhoods during training as well as their performance during repeated testing phases (accuracy and reaction time) to determine the time course of learning these neighborhoods. Subjects sampled the high and low density neighborhoods randomly in early learning, and then over-sampled the high density neighborhood until test performance on both neighborhoods reached asymptote. These results provide a new window on the time-course of learning an artificial lexicon and the role that learners' implicit preferences play in learning highly confusable words.

**Keywords:** spoken word recognition; phonological neighborhoods; word learning; artificial lexicon

## Introduction

Since the pioneering work of Marslen-Wilson (1987) on the role of acoustic/phonetic similarity in on-line spoken word recognition, there has been debate over the structure of phonological neighborhoods in the mental lexicon. In the cohort model, lexical items were neighbors—and, thus, competitors—if and only if their sound-forms overlapped from the beginning of the word, such as in “pat” and “pack”. The Neighborhood Activation Model (Luce, Pisoni & Goldinger, 1991; Luce & Pisoni, 1998) quantified neighborhood similarity as a combination of factors: the frequency of the single item in question, neighborhood density (also describable as confusability), and overall neighborhood frequency. Neighbors as defined by NAM can include rhyme words (e.g. “pat” and “rat”) and words with other one-segment differences, such as “pat” and “past”.

A series of studies by Creel, Aslin, and Tanenhaus (2006) employing an artificial lexicon (Magnuson, Tanenhaus,

Aslin, & Dahan, 2003) further revealed another intricacy of neighborhood structure, specifically by asking whether all segment differences, regardless of type (i.e., consonant vs. vowel) have an equal influence on confusion of newly learned words. Two CVCV items with matching consonants are more often confused with each other than two CVCV items with matching vowels; in other words, segment type matters. Furthermore, the position of the consonants played a role: VCVC items with matched consonants did not elicit such confusions.

One outstanding question concerning neighborhoods is how they develop. After a pre-lexical infant learns its first word (e.g., “no” or its own name), how does acoustic/phonetic similarity affect the learning of new words? Do infants acquire words based solely on frequency of occurrence in the ambient linguistic environment, or do they systematically avoid attending to novel words that are similar in sound-structure with known words? Based on corpus analyses, Charles-Luce and Luce (1990, 1995) made just such a prediction, but others have provided conflicting evidence (Coady & Aslin, 2003; Dollaghan, 1994). More direct evidence comes from word-learning studies with toddlers. Swingley and Aslin (2007) taught young children new words that were either neighbors to words they already knew (e.g., “tog” vs. “dog”) or non-neighbors (e.g., “meb”). Neighbor items were more difficult than non-neighbors for the children to learn. However, conflicting evidence from toddlers exists (Newman, Samuelson & Gupta, 2008), suggesting that with more exposure they can learn a novel item from a high-density neighborhood as well as from a very low-density neighborhood.

This same question of neighborhood effects on word learning applies to adults, who are constantly acquiring new words in their lexicon (e.g., “locavore”, “staycation”). Perhaps more relevant to the growing adult lexicon is the case of learning words in a second language. Here there is both a neighborhood effect *within* the L2 lexicon and interference effects *between* the L1 and L2 lexicons. There is conflicting evidence of between-language neighborhood effects (Spivey & Marian 1999; Ju & Luce, 2004) in spoken word recognition, but virtually no evidence for such effects in word *learning*. One reason for this limited evidence is that studies of L1 and L2 lexicons are extremely difficult to control, and L2 often provides the learner with phonetic and

phonological cues that clearly mark the lexical item as a member of only one language.

Another approach to the study of neighborhood effects in word learning is to create new words that are designed to compete with known words. Gaskell & Dumay (2003) present evidence of competition development in English-speaking adults who learn a non-English word that competes with an English word that lacks neighbors. For example, “cathedral” has no English neighbors, but listeners were exposed to the meaningless word-form “cathedruke” over the course of an experiment. The novel item immediately leads to facilitatory effects on the English item. However, after sleeping, subjects’ behavior reflected lexical competition between the two forms. These results suggest that new words compete with old words during spoken word recognition, but they do not bear directly on the time-course of *learning* new words. Importantly, Magnuson et al. (2003), using an artificial lexicon, found no significant evidence that the native language (English) interfered with the processing of neighbors from the artificial lexicon, at least not after only 2 hours of training. Thus, in the early phase of training, even with 90% or better accuracy in learning the names for novel objects, adults do not seem to show between-language neighborhood interference.

Here we describe a study of adult learners using an artificial lexicon. The rationale for using an artificial lexicon, as in Magnuson et al. (2003) and Creel et al. (2006), is that we can carefully control all the parameters of the lexicon (density, frequency, phoneme inventory, meaning) that are very difficult to balance using natural language materials. Our key innovation is creating a learning paradigm in which adults choose how they listen to the entire set of novel words. They must map 16 novel word-forms onto 16 novel visual shapes. Across a series of learning blocks, subjects sample the sound-object pairs by selecting a shape on a touch screen and hearing that shape’s name. A testing phase after each training block assesses the accuracy and speed of word recognition using the same touch screen. By varying the neighborhood structure within the set of 16 words, we can determine whether adult learners choose to sample from high or low density neighborhoods during the process of learning novel word-object mappings.

### Overview of Design

The learning environment was simplified by presenting each subject with an array of 16 novel shapes on a touch screen display. We selected a touch-screen monitor rather than a computer mouse because it lends itself to ease of use for children or other special populations who have limited mouse experience. Subjects are instructed that they should learn the names of the shapes by touching them and hearing that shape’s name. They are told that they can touch shapes in any order and that they have 64 touches per block. After each block, they will be tested for their knowledge of shape names by hearing a name and completing a 16-AFC task. This alternation of training blocks and testing blocks allows us to describe any changes in how subjects allocate their

exploration of the lexicon and its relation to how well subjects have learned the lexicon.

## Method

### Participants

A total of 41 subjects from the University of Rochester participated in the study and consented per the guidelines of the University of Rochester human subjects review board. Each subject received \$10 for one session of approximately 45 minutes. They were told that they would be listening to words and selecting pictures on a touch screen, to learn the names of the pictures, and subsequently tested on what they have learned. All subjects reported normal hearing, normal or corrected-to-normal vision, and were native speakers of English.

### Stimulus Materials

The lexicon was created to vary in neighborhood density and type of acoustic/phonetic similarity. It consisted of 16 items in total: a high-density cohort neighborhood (baga, bagi, bago, bagu), a high-density rhyme neighborhood (dido, kido, pido, tido) and 8 low-density items (gobu, dupi, poti, toku, kuba, tupa, gota, puki). Items were recorded as WAV files by a graduate student with linguistics training. The speaker read each item at a natural rate, yielding an average word length of 745 milliseconds. Items were paired with 16 novel black and white images (Hunt & Aslin, 2009). Three different list conditions of random pairings of words and pictures were used, counterbalanced to avoid any item effects that may have arisen from particular word-image pairings. Subjects were randomly assigned to conditions.

### Environment

The experiment was run on MathWorks Matlab and the Matlab Psychophysics toolbox (Brainard, 1997; Pelli, 1997) on a Dell Dimension desktop PC running Windows XP with an NEC touchscreen monitor. Images were randomly presented in a 4x4 grid on the screen., as seen in Figure 1. Subjects listened to words over Sennheiser HD 570 headphones set at a comfortable volume level. The study was conducted in a sound-attenuated booth.

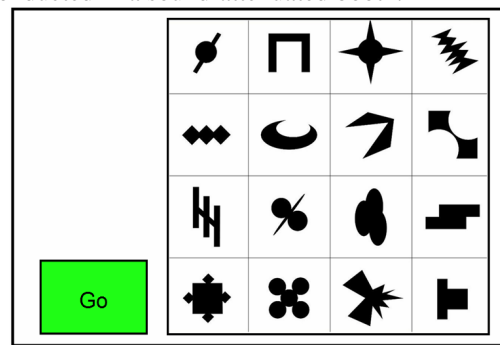


Figure 1: An example screen

## Task

Training and testing were alternated in a session, with 6 blocks of each, for a total of 12 blocks. Participants were told to select the items (i.e., touch a shape) in any order they desired to learn the words that named each of the 16 images on the screen. They were not told that they would be trained on the same 16 images and corresponding words in future blocks. Rather, they were told that each testing block would correspond to the preceding training block. During each training block, an on-screen counter marked off the number of remaining training trials the subject had, from 64 to 0, until a test block would begin. If the subject had evenly distributed their touches in a block, they would hear each word 4 times, which was deemed sufficient for a minimal level of familiarity but not full mastery after the initial block. The location of each image was randomized four times during training blocks: once at the beginning of the block, then once after each 16 trials. This precluded the possibility that subjects made associations between item location and name, rather than the desired effect of item image (shape) and name.

A test block consisted of two passes through the list of lexical items, for a total of 32 trials, in random order. Subjects pressed a GO button image on the screen to start a test trial, then heard a word corresponding to one of the 16 images on the screen. They were free to select any of the items present on the screen, in a 16-AFC task. Instructions specified that they should respond as quickly and accurately as possible. If they correctly selected an image, the image turned green. If they selected an incorrect image, the image turned red; they were not informed of the correct image. Thus feedback provided the subject with only minimal information about each decision—whether it was correct—but not information as to which image was the correct item if they made an error. Allowing the participant to start each trial provided the opportunity for short rests as needed during the testing blocks.

## Results

### Training data

An analysis of variance was performed to determine the effect of density (high versus low) on overall proportion of selection of training items; the result was not significant. Within high density items, a two-factor ANOVA with replication was performed, comparing the number of selections of items from the high density cohort neighborhood to those from the high density rhyme neighborhood, across blocks. Cohort items were chosen more frequently than rhyme items,  $F(1,84) = 15.69, p < .001$  (see Figure 2). Block was also a significant factor  $F(5,420) = 2.52, p < .05$  and there was an interaction of neighborhood type and block,  $F(5, 420) = 3.60, p < .01$ .

Proportion of training selection

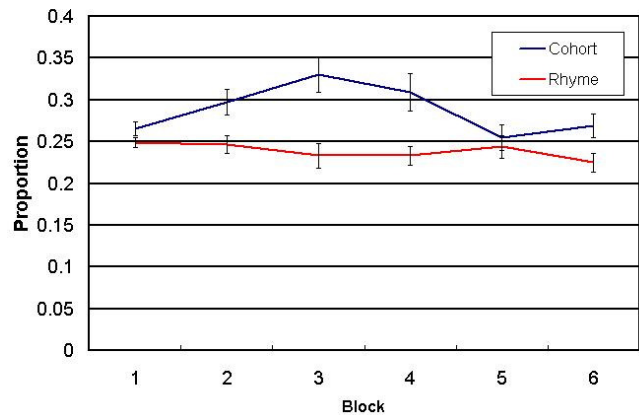


Figure 2: Training selections of high density items

Training sequences were then analyzed for the likelihood that a subject, having selected an item from a particular neighborhood, would next select an item from the same neighborhood. The blue line in Figure 3 shows, across all subjects, the proportion of item selections that, on the immediately following trial, were drawn from the same high density neighborhood. Error bars represent standard errors of the mean. The pink line represents the eight low-density items grouped into random pseudo-neighborhoods of four items, to provide a baseline comparison for the likelihood of selecting within any group of four items.

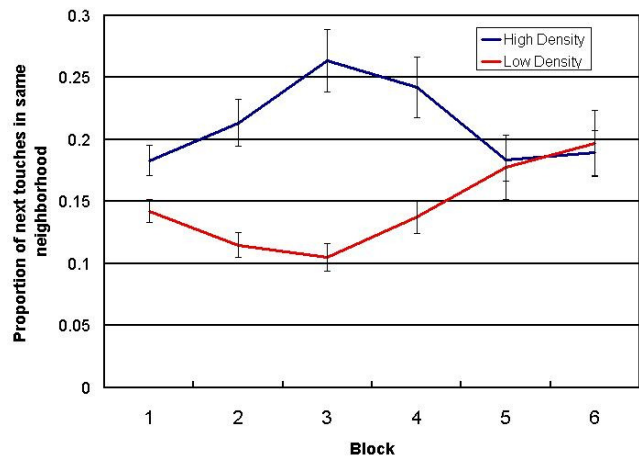


Figure 3: In-neighborhood probability

A two-factor ANOVA with replication was performed. This revealed a main effect of density, low versus high,  $F(1,5) = 37.86, p < .001$ , and a significant interaction of density x block,  $F(1,5) = 5.92, p < .001$ . Within the high-density neighborhoods, a single factor ANOVA examined whether there was an effect of block on probability of successive same-neighborhood selections. A significant effect of block was found,  $F(5,252) = 2.94, p < .05$ .

### Test data:

Figure 4 shows that for both high and low density items, the accuracy of responding on the 16-AFC test blocks rose rapidly from 50% correct (chance=6.25%) to asymptotic performance within the 6 testing blocks. A two-factor ANOVA with replication examining accuracy across test blocks revealed a significant effect of block  $F(5,504) = 98.19, p < .001$ . There was also an effect of density  $F(1,1) = .10, p < .05$ , but no interaction of density with block  $F(5,504) = .36, p > .05$ .

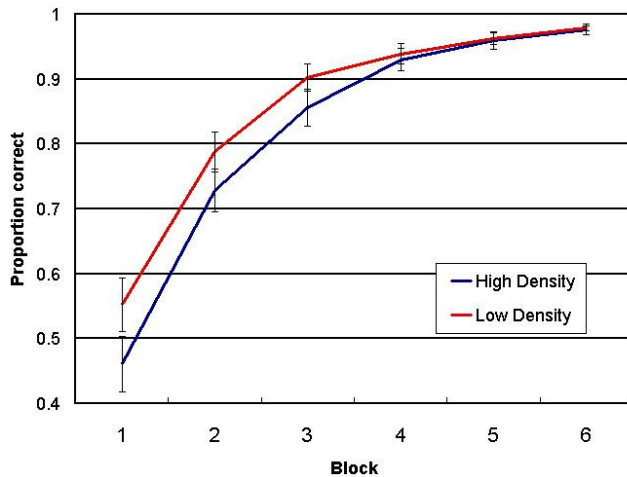


Figure 4: Proportion of test trials correct

A single-factor ANOVA of reaction times for correct trials across test blocks, collapsed across densities, showed a significant effect of block  $F(5,252) = 8.03, p < .001$ .

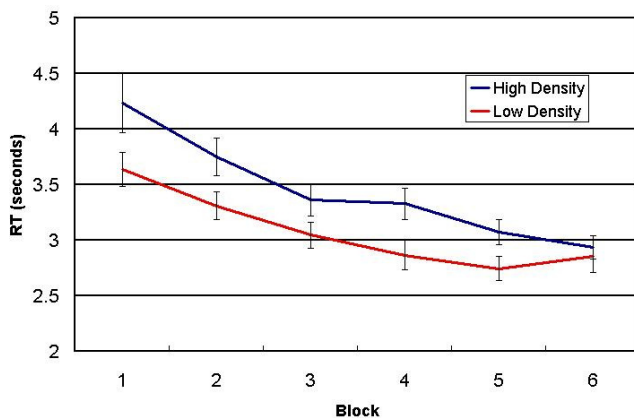


Figure 5: Reaction times in test

Figure 5 shows reaction times in the initial and final test blocks, as a function of neighborhood density. A two-factor ANOVA with replication revealed highly significant effects of block  $F(5,492) = 14.50, p < .001$  and density  $F(1,1) = 18.36, p < .001$ .

## Discussion

The present experiment is the first that we know of to assess how human learners allocate their attention by selecting novel words for association with novel visual objects in a word-learning paradigm. By having subsets of words that share acoustic/phonetic properties (lexical neighbors), we could ask whether learners seek or avoid repetitive samples of words from low- or high-density neighborhoods as they acquire new word-object associations.

### Training data

Throughout the six training blocks, more high density cohort items were selected than high density rhyme items. Subjects concentrated their selections on cohort items, presumably because of the perceived phonological similarity among these items.

For high density items, there was a significant effect of block on the likelihood that subjects selected an item from one neighborhood and then on the subsequent touch selected an item from the same neighborhood. (This includes pressing the same item twice in a row.) In the initial trials, selection was nearly random. As the session continued, however, the probability that subsequent selections were within the same neighborhood significantly increased, then decreased to initial levels as mastery of the items was achieved and concentrated training on neighbors was no longer beneficial. When the low density items were randomly grouped into two groups of four and the selection data from those were compared to the high density selections, there was a significant effect of density and an interaction with blocks. Subjects were more likely to select two high-density items within a neighborhood (of four), one after another, than to select any two low-density items out of a random grouping of four. A regression model (proposed later) may be informative in further analyses of the influence of word-sampling behavior within one training block on the behavior exhibited in subsequent blocks.

### Testing data

Subjects achieved 51% accuracy within the first testing block, after hearing each word on average only 4 times; this is significantly above chance (6.25% correct). This minimal exposure was sufficient to achieve significant learning, but performance did not reach asymptotic levels until 5 or 6 blocks of training. Accuracy was affected by density. High density items were correctly identified less frequently than low density items until halfway through the experiment. Their phonological similarity presumably created greater difficulty for the subjects.

As in previous studies, differences in reaction times also occurred as a result of density, with low density items being responded to more rapidly than high density items consistently until the final block, at which point performance on the two densities was equivalent.

## Future studies

Numerous ways of examining the development of neighborhoods could provide greater insight into the time course of lexical competition during word learning. One such study would be to successively reveal subsets of neighborhoods to the learner in the paradigm described here. The training sequences would be of particular interest; in contrast to the present study, a different set of items would be present during each training block, and so the subject may adjust training strategies accordingly as overall neighborhood density is revealed across blocks.

Other statistical analyses, in the form of linear regression models, may reveal more about the present study and future designs. One analysis would be whether the performance on one test block influences training patterns in the immediately following block, which the current analyses cannot address well. Finally, our current analyses of within-neighborhood effects of training include pairs of trials in which one of the four items from the same high-density neighborhood is selected, including immediate repeats of the same item. Excluding these identical repeats may be more relevant to the question of lexical competition during learning.

## Conclusion

It is well established that words in high-density neighborhoods are more difficult to process than words in low-density neighborhoods. When adults are given control over the frequency of exposure to novel words, they quickly adjust the rate of exposure by over-sampling words in high-density neighborhoods, particularly cohort neighbors more so than rhyme neighbors. The difficulty of learning high density items was revealed in this study as differences in accuracy and reaction time, which persist for the initial blocks of test trials. However, subsequent training yields equivalent learning accuracy for words in high- and low-density neighborhoods.

Our paradigm is likely to be useful for addressing a variety of issues in lexical learning. Perhaps most importantly, the method may be useful for teaching children novel lexical items, either in the laboratory or in the classroom.

## Acknowledgments

This research was supported by NIH Grants HD037082 to RNA and DC005071 to MKT. The authors express gratitude to Johnny Wen for programming assistance and Dana Subik and Carrie Miller for recruiting participants.

## References

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433-436.

Charles-Luce, J. & Luce, P. A. (1990). Similarity neighbourhoods of words in young children's lexicons. *Journal of Child Language 17*, 205-15.

Charles-Luce, J. & Luce, P. A. (1995). An examination of

similarity neighbourhoods in young children's receptive vocabularies. *Journal of Child Language 22*, 727-35.

Coady, J. A., & Aslin, R. N. (2003). Phonological neighbourhoods in the developing lexicon. *Journal of Child language, 30*(02), 441-469.

Creel, S. C., Aslin, R. N., & Tanenhaus, M. K. (2006). Acquiring an artificial lexicon: Segment type and order information in early lexical entries. *Journal of Memory and Language, 54*, 1-19.

Dollaghan, C. A. (1994). Children's phonological neighbourhoods: half empty or half full? *Journal of Child Language 21*, 257-71.

Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition, 89*(2), 105-132.

Hunt, R. H., & Aslin, R. N. (2009). Category induction via distributional analysis: Evidence from a serial reaction time task. *Journal of Memory and Language, 62*, 98-112.

Ju, M. & Luce, P. A. (2004). Falling on Sensitive Ears - Constraints on Bilingual Lexical Activation. *Psychological Science, 15* (5), 314-318.

Luce, P. A., Pisoni, D. B., & Goldinger, S. D. (1991). Similarity neighborhoods of spoken words. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing* (pp. 122-147). Cambridge, MA: MIT Press.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19*(1), 1-36.

Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N., & Dahan, D. (2003). The time course of spoken word learning and recognition: Studies with artificial lexicons. *Journal of Experimental Psychology: General, 132*(2), 202-227.

Newman, R., Samuelson, L., & Gupta, P. (2008). *Learning Novel Neighbors: Distributed mappings help children and connectionist models*. Paper presented at the Cognitive Science Society, Washington, DC, USA.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437-442.

Revill, K. P., Tanenhaus, M. K., & Aslin, R. N. (2008). Context and spoken word recognition in a novel lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(5), 1207-1223.

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science, 10*(3), 281-284.

Swingle, D., & Aslin, R. N. (2007). Lexical competition in young children's word learning. *Cognitive Psychology, 54*(2), 99-132.