Effects of Near and Distant Phonological Neighbors on Picture Naming

Daniel Mirman (mirmand@einstein.edu)
Moss Rehabilitation Research Institute
1200 W. Tabor Rd., Philadelphia, PA 19141, USA

Audrey K. Kittredge (akittre2@illinois.edu)
Gary S. Dell (gdell@cyrus.psych.uiuc.edu)
Beckman Institute, University of Illinois
405 N. Matthews Ave., Urbana, IL 61801, USA

Abstract
Many studies have examined the effects of co-activation of similar words (“neighbors”) during processing, with some reporting facilitative effects and others reporting inhibitory effects. Attractor dynamics has provided a promising integrated account in which distant semantic neighbors (moderately similar words) tend to facilitate processing and near semantic neighbors (highly similar words) tend to inhibit processing. This framework was extended to phonological neighbor effects on the accuracy of word production. For aphasic patients (N=62) and speeded young controls (N=32), picture naming was more accurate for words with many distant phonological neighbors (words with matching onsets) and less accurate for words with a near phonological neighbor (homophones). In addition, the sizes of the facilitative and inhibitory effects were correlated, suggesting that the mechanisms responsible for both effects are functionally integrated. These results extend an attractor dynamics framework that predicts facilitative effects of distant neighbors and inhibitory effects of near neighbors.

Keywords: phonological neighbors; cohort density; homophones; neighborhood density; word production; attractor dynamics.

Introduction
Theories of language processing agree that similar words are co-activated during processing. Such co-activation provides a simple account of classic priming effects: processing cat partially activates dog (due to semantic similarity) and can (due to form similarity), facilitating responses to those words (Marslen-Wilson & Zwitserlood, 1989; Meyer & Schvaneveldt, 1971; Zwitserlood, 1996). Co-activation is also consistent with findings from studies using the visual-world paradigm: when instructed to click on a picture of a cat, participants are more likely to fixate images of a dog or a can (Allopenna, Magnuson, & Tanenhaus, 1998; Huettig & Altmann, 2005; Magnuson, Tanenhaus, Aslin, & Dahan, 2003; Mirman & Magnuson, 2009; Yee & Sedivy, 2006). Co-activation of similar words has also been used to account for global similarity effects: the number of similar words that are likely to be co-activated given a particular similarity metric, called “neighborhood density”. In this report, we examine the effects of two kinds of phonological neighbors on word production in aphasic patients and speeded young controls.

The effects of neighborhood density on word processing are complex and poorly understood. Neighbors defined by form similarity (spelling or sound) have been found to facilitate printed word recognition (e.g., Sears, Hino, & Lupker, 1995) and spoken word production (e.g. Vitevitch, 2002). However, phonologically similar neighbors consistently produce inhibitory effects in many tasks involving spoken word recognition (e.g., Luce & Pisoni, 1998; Magnuson, Dixon, Tanenhaus, & Aslin, 2007). Neighbors defined by semantic similarity can also exert effects in both directions. Near neighbors (concepts with very similar meanings) inhibit word recognition and distant neighbors (concepts with moderately similar meanings) facilitate visual word recognition (Mirman & Magnuson, 2008). Mirman and Magnuson suggested that this contrast between the impact of near and distant neighbors on word processing may be a general property of word processing. For example, although orthographic neighbors (salt - halt) generally facilitate visual word recognition, transposed-letter neighbors (salt - slat) exert inhibitory effects (Andrews, 1996).

The attractor dynamics framework for cognition represents each concept as a stable state (“attractor basin” or simply “attractor”) in a high-dimensional space of possible mental states (for an accessible introduction see Spivey, 2007). Word processing is a matter of traversing this space in order to reach the correct attractor. When the system has reached a stable state, it is deemed to have “settled” and the accuracy of the system’s final state can be compared relative to the target attractor. Neighbors are other attractors and distance between attractors is determined by similarity. The critical insight from attractor dynamics is that different similarity relations between neighbors can exert different effects on the settling process (Mirman & Magnuson, 2008). Distant neighbors create a broader attractor basin, which facilitates settling to the correct attractor. In contrast, near neighbors are too few to substantially change the overall size of the attractor basin, but because of their high similarity (i.e., proximity) to the target, they function as conflicting subbasins, which slows the completion of the settling process.

An alternative to the attractor dynamics account would be to simply stipulate that neighbor effects are different in
different contexts or tasks. For example, Vitevitch and Luce (1998; 1999; see also Luce & Large, 2001) proposed that, in speech perception, sub-lexical neighbor effects are facilitative and lexical neighbor effects are inhibitory. However, there are three arguments against this view as a general account of neighborhood effects. First, the empirical data have been challenged (Lipinski & Gupta, 2005). Second, semantic neighbor effects appear to emerge at a single level of processing (i.e., semantics), thus, assigning different effects to different levels cannot account for the facilitative effects of distant semantic neighbors and the inhibitory effects of near semantic neighbors (Mirman & Magnuson, 2008). Third, it is unparsimonious to propose that neighbor interactions have fundamentally different properties at different levels of processing.

Attractor dynamics provide a parsimonious, integrated account in which neighbors can have different, context-dependent effects. However, the existing data have only examined the key attractor dynamics prediction in the domain of semantic neighborhoods. The present studies examine these same predictions in the domain of phonological neighbor effects on word production.

As noted above, previous studies have found facilitative effects of phonological neighbors on word production. Vitevitch (2002) found that healthy young controls produced more errors in an error-elicitation paradigm and were slower to name pictures for words with few phonological neighbors compared to words with many phonological neighbors. Similarly, aphasic patients produce more errors when naming pictures with low phonological neighborhood density names (Gordon, 2002; Kittredge, Dell, Verkuilen, & Schwartz, 2008).

Given the facilitative effect of phonological neighbors in picture naming tasks, one might expect that greater phonological similarity would strengthen this effect. In the extreme case, words with different meanings but identical phonological forms, that is, homophones (e.g. can [container] vs. can [able]) might be particularly easy. After all, the naming target’s homophone is maximally phonologically similar to target’s phonology. However, if both meanings are activated during an attempt to retrieve the name of one meaning of a homophone, those meanings may compete, consequently producing slower responses and higher error rates. Thus, there is reason to expect the opposite result. Indeed, this is the critical prediction from the attractor dynamics account of neighborhood effects.

There is an extensive experimental literature investigating homophony in word production, much of which is concerned with whether word frequency effects on picture naming arise from syntactic-semantic representations or phonological form representations (e.g. Caramazza, Costa, Miozzo, & Bi, 2001; Jescheniak & Levelt, 1994). Although there is no consensus among these studies, it is likely that production latencies for a homophone are influenced by the frequency of both its meaning and its form. More relevant to our analysis are findings that both meanings of a homophone are activated during word production. For example, priming the non-pictured homophone meaning affects response latency and accuracy in picture naming tasks (Cutting & Ferreira 1999; Ferreira & Griffin, 2003). Moreover, picture naming studies with aphasic patients have shown that training on one homophone meaning generalizes to the other meaning (Biedermann, Blanken, & Nickels, 2002; Biedermann & Nickels, 2008a; Biedermann & Nickels, 2008b).

These studies suggest that homophone production involves some degree of interaction between the target and its homophone mate. Given this, if homophones are viewed as very near phonological neighbors, the attractor dynamic approach of Mirman and Magnuson (2008) predicts that having a homophone should be associated with some kind of cost. Alternately, if the extreme similarity of the homophone just exaggerates the positive effect of having a similar neighbor, then the expectation is for a benefit. These conflicting predictions were tested by examining the accuracy of picture naming in aphasic patients and in speeded young controls.

### Experiment

**Methods**

**Participants.** There were two sets of participants: aphasic patients and speeded young controls. The patients were 62 unselected aphasic patients recruited from the MRRI Cognitive Rehabilitation Research Registry (Schwartz, Brecher, White, & Klein, 2005) on the basis of chronic aphasia secondary to left cerebral vascular accident. They had a mean age of 58 (range 26–78), mean years of education of 14 (10–21), and most (over 90%) were at least 6 months post-onset. The patients were all premorbidly right-handed, had English as the primary language, adequate vision and hearing, and unilateral left hemisphere damage (not restricted to subcortical areas). These patients included all aphasia subtypes and covered a wide range of performance (2%–97% correct naming). The young controls were 32 healthy college students with no known history of neurological, visual, or auditory impairments, who were recruited from the University of Illinois participant pool.

**Materials.** The 175-item Philadelphia Naming Test (PNT; Roach et al., 1996) was used to measure word production in picture naming. The black and white pictures represent objects from varied semantic categories and have high familiarity, name agreement, and image quality. Names range in length from 1 to 4 syllables and in frequency (normalized to occurrences per 1 million word tokens) from 1 to 100.

Our concern is with the effects of “near” and “distant” phonological neighbors on picture naming in order to test...
the general attractor dynamics prediction that the effect of neighbors will depend on their impact on the attractor landscape. Distant phonological neighbors were defined as words that share onsets with the target word. These words are described as “cohorts” because they form the cohort of partially activated words during spoken word recognition (e.g., Allopenna et al., 1998; Magnuson et al., 2007; Marslen-Wilson & Zwitserlood, 1989). There are many possible phonological neighborhood measures, which are all strongly correlated with one another. The cohort density measure (the summed log frequency of the target word and all of its cohorts) was chosen because word onsets are particularly important for spoken word processing.

Lexical variables (phonological neighborhood, word frequency, etc.) were assessed using the American National Corpus (Ide & Suderman, 2004), a large-scale, representative corpus of American English containing over 3.2 million spoken word tokens. The words in the PNT were divided into “few” and “many” neighbor conditions based on the median cohort density (31.5) and a few words were eliminated to ensure that the conditions had an equal number of words and were matched in word frequency and length (the resulting conditions were composed of 85 words each). Table 1 shows that the two conditions were matched in word frequency and length and strongly different in cohort density as well as differing on other phonological neighborhood measures. For the purpose of this experiment, it was not necessary to isolate a particular measure of phonological neighborhood; rather, it was sufficient that words in the two conditions strongly differed in their number of phonologically similar words.

Table 1. Mean (standard deviations in parentheses) properties of stimuli for cohort density manipulation.

<table>
<thead>
<tr>
<th></th>
<th>Few neighbors</th>
<th>Many neighbors</th>
<th>t</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Phonological neighborhood measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort Density</td>
<td>14.9 (8.7)</td>
<td>73.2 (36.0)</td>
<td>14.5</td>
<td>0.0001</td>
</tr>
<tr>
<td>Neighborhood Density</td>
<td>10.2 (8.8)</td>
<td>14.1 (11.3)</td>
<td>2.5</td>
<td>0.05</td>
</tr>
<tr>
<td>Number of Neighbors</td>
<td>11.8 (12.8)</td>
<td>16.2 (15.2)</td>
<td>2.0</td>
<td>0.05</td>
</tr>
<tr>
<td>Posit. Prob.</td>
<td>.211 (0.1)</td>
<td>.263 (0.1)</td>
<td>3.2</td>
<td>0.01</td>
</tr>
<tr>
<td>Transit. Prob.</td>
<td>.017 (0.02)</td>
<td>.023 (0.02)</td>
<td>2.3</td>
<td>0.05</td>
</tr>
<tr>
<td><strong>Control Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Num. Words</td>
<td>85</td>
<td>85</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Log Frequency</td>
<td>1.07 (0.7)</td>
<td>1.16 (0.7)</td>
<td>0.94</td>
<td>0.35</td>
</tr>
<tr>
<td>Num. Letters</td>
<td>5.51 (1.9)</td>
<td>5.11 (1.9)</td>
<td>1.4</td>
<td>0.17</td>
</tr>
<tr>
<td>Num. Phonemes</td>
<td>4.33 (1.7)</td>
<td>4.35 (1.5)</td>
<td>0.09</td>
<td>0.93</td>
</tr>
</tbody>
</table>

Near phonological neighbors were defined as words with identical phonological forms and unrelated meanings, that is, homophones. The 175 words in the PNT include 14 homophones for which the pictured meaning is the dominant meaning (meaning dominance was assessed based on proportion of associated words in the USF free association norms (Nelson, McEvoy, & Schreiber, 2004): $M=73.6\%$, $SD=10.5$, Range=50.4-86.7). Number of meanings (homophony) was assessed based on the number of distinct entries in the online Wordsmyth dictionary (http://new.wordsmyth.net/). For each of these homophones a control (unambiguous) word was selected from the PNT that was matched to the homophone on word frequency, length, and phonological neighborhood variables (see Table 2).

Table 2. Mean (standard deviations in parentheses) properties of stimuli for homophony manipulation.

<table>
<thead>
<tr>
<th></th>
<th>Homophones</th>
<th>Control Words</th>
<th>t</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Num. Meanings</td>
<td>2.21 (0.58)</td>
<td>1.0 (0)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Control Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Num. Words</td>
<td>14</td>
<td>14</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cohort Density</td>
<td>50.6 (41.1)</td>
<td>46.7 (36.0)</td>
<td>0.87</td>
<td>0.40</td>
</tr>
<tr>
<td>Neighborhood Density</td>
<td>26.1 (14.7)</td>
<td>27.2 (15.2)</td>
<td>0.30</td>
<td>0.77</td>
</tr>
<tr>
<td>Number of Neighbors</td>
<td>22.4 (9.8)</td>
<td>22.00 (9.1)</td>
<td>0.22</td>
<td>0.83</td>
</tr>
<tr>
<td>Posit. Prob.</td>
<td>.202 (0.04)</td>
<td>.195 (0.06)</td>
<td>0.55</td>
<td>0.59</td>
</tr>
<tr>
<td>Transit. Prob.</td>
<td>.014 (0.01)</td>
<td>.013 (0.01)</td>
<td>0.61</td>
<td>0.55</td>
</tr>
<tr>
<td>Log Frequency</td>
<td>1.47 (0.80)</td>
<td>1.40 (0.50)</td>
<td>0.54</td>
<td>0.60</td>
</tr>
<tr>
<td>Num. Letters</td>
<td>4.07 (0.92)</td>
<td>4.00 (0.78)</td>
<td>1.00</td>
<td>0.34</td>
</tr>
<tr>
<td>Num. Phonemes</td>
<td>3.29 (0.61)</td>
<td>3.29 (0.61)</td>
<td>0.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Procedure. The patients were tested using the standard PNT procedure (http://www.ncrrn.org/assessment/pnt; Roach et al., 1996; see also Dell et al., 1997; Schwartz et al., 2006): each picture was presented one at a time and the first complete (i.e. non-fragment) response produced within 20 s was scored. The young controls were tested using the tempo picture naming procedure (Hodgson & Lambon Ralph, 2008). This task provides a valuable source of converging data for comparison with the patient data because it has been shown to induce some characteristic aspects of aphasic picture naming errors. In the tempo picture naming task, participants heard a series of beeps set to a tempo (500 ms). On the fourth beep they were also presented with a picture (one of the PNT items), which they were to name and to time their response to coincide with the fifth beep.

Results

Cohort Density. The left panel of Figure 1 shows that picture naming accuracy was lower for low cohort density words than for high cohort density words (Patients: 66.7% vs. 70.4%, $t(61)=5.5$, $p<0.0001$; Speeded controls: 79.7% vs. 81.7%, $t(31)=2.37$, $p=0.05$). Patients also produced more nonword errors for low cohort density words than high cohort density words (8.24% vs. 6.62%, $t(61)=3.23$, $p<0.01$). Speeded controls produced very few nonword
errors ($M=0.68\%, \ SD=0.89\%$) and the numerical trend in the same direction as the patients (0.77\% vs. 0.63\%) was not significant ($t(31)=0.49, \ p>0.6$). There were no significant effects on any other error type. The cohort density finding is consistent with previous findings that words with many phonologically similar words are easier to produce (Gordon, 2002; Kittredge et al., 2008; Vitevitch, 2002).

Figure 1. Picture naming accuracy for high and low cohort density words (left panel) and for homophones and control words (right panel). Error bars reflect 1SE.

Homophony. The right panel of Figure 1 shows that participants were more accurate for the control words than for the homophones (Patients: 77.0\% vs 71.7\%, $t(61)=3.45, \ p<0.001$; Speeded controls: 83.5\% vs. 79.4\%, $t(31)=5.43, \ p<0.0001$). This finding is consistent with previous results that indicate slowed processing due to competition between different meanings of homophones (e.g., Shatzman & Schiller, 2004; see also Ferreira & Griffin, 2003). The increased errors for homophones did not aggregate to a specific error type (i.e., no reliable differences for any error type).

Relation between effect sizes. We tested the correlation between cohort density and homophony effect sizes across participants to examine whether there is a possible relationship between them. Figure 2 shows each participant’s homophony effect size (homophones – control) plotted against that participant’s cohort density effect size (high – low). The effect sizes were reliably correlated for patients ($r = -0.25, \ p<0.05$) and for speeded controls ($r = -0.76, \ p<0.0001$).

One possible explanation for this effect size correlation is that there is simply an effect of overall accuracy. That is, participants who make more errors show bigger differences between any conditions. To test this hypothesis, we examined correlations between overall accuracy for the critical conditions and the effect size. For patients, neither correlation approached significance (homophony: $r = 0.0032, \ p > 0.98$; cohort density: $r = 0.0937, \ p > 0.46$). The same was true for controls (homophony: $r = 0.1124, \ p > 0.54$; cohort density: $r = -0.2616, \ p > 0.14$). Since it is not due to overall accuracy, the correlation between cohort density and homophony effect sizes suggests that the mechanisms involved in producing the benefit of similar-sounding words (cohort density effect) are closely tied to those involved in producing the cost of identical-sounding words (homophony effect).

Figure 2. Relationship between homophony and cohort density effect sizes. Open circles correspond to patients, filled triangles correspond to speeded controls.

General Discussion

We examined the effects of phonological neighbors on picture naming in aphasic patients and speeded young controls. Two kinds of phonological neighbors were considered: similar-sounding words defined as words with matching onsets (i.e., cohorts) and identical-sounding words (i.e., homophones). These different phonological neighbor types capture the important distinction between distant and near neighbors. Mirman and Magnuson (2008) found that distant semantic neighbors facilitated word recognition and near semantic neighbors inhibited word recognition. Andrews (1996) found a similar contrast between the effects of (distant) orthographic neighbors and (near) transposed-letter neighbors on visual word recognition. Based on these results, we predicted facilitative effects of phonological neighbors and inhibitory effects of homophony.

The results were consistent with these predictions: both participant groups exhibited a facilitative effect of cohort density and an inhibitory effect of homophony. In addition, the effect sizes were correlated across participants; that is, participants who showed larger cohort density advantage effects also showed larger homophony disadvantage effects. This suggests that the mechanism or mechanisms that produce these effects are functionally integrated.

To account for the contrasting effects of near and distant semantic neighbors, Mirman and Magnuson (2008) proposed an account based on attractor dynamics. On this view, distant neighbors create a broader attractor basin,
which facilitates settling to the correct attractor. In contrast, near neighbors are too few to substantially change the overall size of the attractor basin, but because of their high similarity (i.e., proximity) to the target, they function as conflicting subbasins, which slows the completion of the settling process. These distinctions are shown schematically in Figure 3. Mirman and Magnuson confirmed this account using simulations of a computational model.

Figure 3. Top: Schematic diagram of narrow and broad attractor basins resulting from few and many distant neighbors, respectively. Bottom: Schematic diagram of a single attractor basin and an attractor with a subbasin formed by a near neighbor.

To extend this framework to word production it is helpful to consider picture naming as a process of settling to an attractor in a multidimensional space that combines semantic and phonological dimensions. For a given target word, cohort neighbors and homophone neighbors are equally semantically unrelated to the target (the cohort pair can – cat and the homophone pair can [container] – can [able] are equally semantically unrelated). On the phonological dimensions, the homophone neighbors have substantially more similarity to the target word than cohort neighbors do (i.e., complete phonological overlap vs. shared onsets). Therefore, a large number of cohort neighbors can increase the gradient and facilitate settling to the correct attractor. In contrast, a single homophone neighbor will not have a substantial impact on the gradient, but can form a competing subbasin, which can delay the settling process.

If the settling process is disrupted by damage or a time constraint, the system may fail to settle completely (no response) or may settle to an incorrect attractor (error). Since settling is facilitated by distant (cohort) neighbors and inhibited by near (homophone) neighbors, this account captures the observed pattern of facilitative effects of cohort density and inhibitory effects of homophony. The correlation between effect sizes could reflect the average sharpness of attractor basins in the landscape. In a landscape with relatively sharp attractor basins, distant neighbor attractors would have a relatively small impact on slope steepness and near neighbors would be less likely to act as a competing subbasin. Attractor dynamic models generally develop sharper attractors over the course of learning, so this individual difference variable could reflect language skill. Further research is required to test this hypothesis or other possible explanations of the correlation between effect sizes.

In sum, the present results demonstrate contrasting effects of near and distant phonological neighbors on picture naming and provide a new perspective on the mechanisms involved in word production. Furthermore, they contribute to the creation of a unified theory of neighborhood effects in lexical processing.

Acknowledgements

This research was supported by the Moss Rehabilitation Research Institute and National Institutes of Health grants DC000191 and HD44458. We thank Myrna Schwartz for her thoughtful suggestions.

References


