Orthographic and phonological neighborhood effects in handwritten word perception

Anthony S. Barnhart¹ · Stephen D. Goldinger²

Published online: 26 August 2015
© Psychonomic Society, Inc. 2015

Abstract In printed-word perception, the orthographic neighborhood effect (i.e., faster recognition of words with more neighbors) has considerable theoretical importance, because it implicates great interactivity in lexical access. Mulatti, Reynolds, and Besner *Journal of Experimental Psychology: Human Perception and Performance*, 32, 799–810 (2006) questioned the validity of orthographic neighborhood effects, suggesting that they reflect a confound with phonological neighborhood density. They reported that, when phonological density is controlled, orthographic neighborhood effects vanish. Conversely, phonological neighborhood effects were still evident even when controlling for orthographic neighborhood density. The present study was a replication and extension of Mulatti et al. (2006), with words presented in four different formats (computer-generated print and cursive, and handwritten print and cursive). The results from Mulatti et al. (2006) were replicated with computer-generated stimuli, but were reversed with natural stimuli. These results suggest that, when ambiguity is introduced at the level of individual letters, top-down influences from lexical neighbors are increased.

Keywords Word perception · Orthography · Phonology · Visual word recognition

Word recognition in both the visual and auditory modalities occurs through a complex pattern of competition and facilitation between entries in the lexicon (e.g., Ziegler & Muneaux, 2007). As such, *neighborhood density* is a variable of key theoretical interest. Coltheart, Davelaar, Jonasson, and Besner (1977) quantified lexical neighborhoods using the “N metric,” which begins with some target word and sums all potential words that can be formed by changing one letter while maintaining the letter positions (see Yarkoni, Balota, & Yap, 2008, for a more continuous metric). By means of this procedure, LACE, MAKE, and LIKE are neighbors of LAKE. In word-naming tasks, words with many neighbors are typically processed faster than words with few neighbors (Andrews, 1989; Huntsman & Lima, 2002; Sears, Hino, & Lupker, 1995). Andrews (1989; Sears et al., 1995) explained neighborhood effects via the McClelland and Rumelhart (1981) interactive-activation (I-A) model. In this model, upon word onset, all of the sublexical units (i.e., features) of the presented word, many of which are common to its neighbors, become activated. This activation feeds forward, activating letter units that, in turn, activate word units. Activation occurs gradually, in a cascaded manner (McClelland, 1979), such that partially activated letters begin activating words. Meanwhile, activated words return feedback activation to their constituent letters, which then reinforce their associated words in an interactive loop. Recognition is achieved when the activation for one word exceeds its competitors by some criterial degree. Depending on the parameter settings, when a letter string is consistent with many words (dense neighborhood), it receives more feedback during processing, allowing activity at the letter level to reach asymptote more quickly.

Although orthographic neighborhood effects have been frequently replicated, do they truly reflect orthographic processing? Mulatti, Besner, and Job (2003) analyzed all words in the CELEX database and noted that orthographic and

✉ Anthony S. Barnhart
abarnhart@carthage.edu

¹ Department of Psychological Science, Carthage College, 2001 Alford Park Drive, Kenosha, WI 53140, USA

² Department of Psychology, Arizona State University, Tempe, AZ, USA

phonological neighborhood sizes are highly correlated. Thus, it is unclear which factor (phonological or orthographic neighborhood) truly drives neighborhood effects. Mulatti, Reynolds, and Besner (2006) conducted two word-naming experiments to separately assess each variable. Their first experiment assessed phonological neighborhood effects using a stimulus set that controlled orthographic neighborhood density (and many other variables). Under these conditions, a reliable neighborhood effect would call into question previous studies of orthographic neighborhood density that failed to control for phonological neighborhood size. Indeed, Mulatti et al. (2006) showed a reliable 13-ms phonological neighborhood effect. Alone, this finding does not rule out potential orthographic neighborhood effects, independent of phonological effects, so in their second experiment they used an identical methodology, testing for orthographic neighborhood effects while controlling for phonological neighborhood size. Under these conditions, no orthographic neighborhood effect was observed (the difference was only 2 ms).

The present study was largely motivated by Mulatti et al.'s (2006) null finding for orthographic density. Our working hypothesis was that orthographic neighborhood effects may be strongly linked to the visual qualities of the words themselves. Considered from the perspective of an interactive model, orthographic neighborhood effects should be most evident in experimental conditions with degraded visual input. If the stimuli all shared pristine visual features, letter-level processes would quickly resolve toward asymptote, and word-level feedback would have little opportunity to improve performance. Although studies of spoken word recognition have almost exclusively employed natural (i.e., human-generated) stimuli, studies of printed-word perception nearly always involve ambiguity-free, normalized typefaces. As was noted by Manso de Zuniga, Humphreys, and Evett (1991), the universal use of synthetic print (in reading research) may yield an incomplete picture of human word perception, systematically minimizing potential higher-order perceptual capabilities. We may gain a deeper appreciation of such processes by examining perception of handwritten words. In two previous studies that examined lexical effects with handwritten stimuli (Barnhart & Goldinger, 2010; Manso de Zuniga et al., 1991), “top-down” effects on perception were increased, relative to the same effects with printed words. Indeed, some effects that rarely appear for printed text became strongly significant for handwritten text. Manso de Zuniga et al. presented participants with alternating lexical decision and naming trials, using low- and high-frequency words in computer-generated print and human cursive. They found consistently larger frequency effects for cursive items. In an extension of their research, Barnhart and Goldinger (2010) found stronger effects of frequency, regularity, bidirectional consistency, and semantic imageability for cursive words, relative to print. Although frequency and stimulus degradation effects are often additive

(e.g., Becker & Killian, 1977), Barnhart and Goldinger suggested that handwriting qualitatively differs from visual noise or contrast reduction: In handwriting, lexical context can directly disambiguate letters, making top-down processing more important.

In the present study, we reexamined the findings of Mulatti et al. (2006) in a context wherein orthographic effects may be more likely to emerge. In a model such as I-A (McClelland & Rumelhart, 1981), if letter strings were perfectly legible, feedback from words to letters would have little effect on recognition. Handwriting presents a more interesting case. As was discussed by Barnhart and Goldinger (2010), handwritten letters are often ambiguous: The same nominal characters change their physical forms across contexts, and very similar forms may signal different intended characters across contexts. When people write in cursive, their letters often connect together, creating potential problems in segmentation. Given these facts, word-level feedback may be far more important when people perceive handwritten words. Thus, words from dense neighborhoods might provide more exemplars to help resolve the target word, speeding processing.

Our predictions for the phonological neighborhood manipulation were less clear. Two potential outcomes seem equivalently motivated: Phonological neighborhood effects could stay constant across levels of physical ambiguity, or they could be reduced because phonologic activation is hampered by a slowdown on the front end of processing. Mulatti et al. (2006) noted that none of the models they examined could effectively simulate their reported outcomes. However, Perry, Ziegler, and Zorzi (2007) later reported that their CDP+ model was capable of producing both the significant phonological neighborhood effect and the null orthographic neighborhood effect, using the stimuli from Mulatti et al. (2006). The lowest levels of the CDP+ architecture constitute an I-A framework (McClelland & Rumelhart, 1981) that immediately interfaces with an orthographic lexicon. Presumably, orthographic neighborhood effects (if observed) would result from feedback between the orthographic lexicon and letter nodes (Andrews, 1989). Phonological neighborhood effects would likely result from interacting components farther down the processing stream. As units in the orthographic lexicon become activated, activation would feed forward to entries in the phonological lexicon. With a pristine input, items from dense phonological neighborhoods would have an advantage over those from sparse neighborhoods in the selection of appropriate phonological outputs. However, with an input that leads to a slow accumulation of information across features, letters, and entries in the orthographic lexicon, any bootstrapping of activation by dense phonological neighborhoods might be comparatively small.

We replicated and extended Mulatti et al.'s (2006) experiments using stimuli that theoretically traversed a continuum of physical ambiguity: Separate groups of participants saw words in computer-generated print, computer-generated cursive, natural print, and natural cursive. Computer print represented a direct replication, using stimuli that were uniform and segmented (and visually familiar). Computer cursive is uniform (i.e., every instance of a letter is identical), but it is less familiar and nonsegmented. Human print has generally separated letters but lacks uniformity, introducing physical degradation that obscures the letter-level signal within the word-level stimulus. Finally, human cursive is both nonsegmented and nonuniform, complicating signal detection even further. We expected that increasing the physical ambiguity of letters would reverse the experimental outcomes observed by Mulatti et al. (2006). Whereas their data showed a modest phonological neighborhood effect and no orthographic neighborhood effect, we expected handwriting (regardless of style) to elicit robust orthographic effects, stronger than any potential phonological effects.

Experiment 1: Phonological neighborhoods

In Experiment 1, we examined the effects of phonological neighborhood density on naming, using the word list from Experiment 1 of Mulatti et al. (2006). We expected to replicate their findings in the computer-generated print condition, although we note that their original effect was quite small (13 ms). Assuming that phonological feedback would remain relatively constant across script styles, we anticipated one of two outcomes: Either no systematic changes in the effect would emerge across conditions, or there would be a slight reduction in the effect as orthographic-to-phonologic activation became complicated by physical ambiguity.

Method

Subjects Experiment 1 included 160 Arizona State University students who received course credit. Of these volunteers, 37 were randomly assigned to the computer print condition, 42 to computer cursive, 41 to natural print, and 40 to natural cursive. All of the subjects were native speakers of English with self-reported normal or corrected-to-normal vision.

Stimuli The stimuli from Experiment 1 of Mulatti et al. (2006) were generated in four script forms (see Fig. 1). The computer print words were presented in 45-point Courier New font, the computer cursive words in the Bickley Script font, and the natural print and cursive stimuli were collected from two writers (whose handwriting styles were consistently print or cursive) using a Logitech io2 digital pen. The writing

Computer Print	Computer Cursive	Natural Print	Natural Cursive
class	<i>class</i>	class	<i>class</i>
gait	<i>gait</i>	gait	<i>gait</i>
hoax	<i>hoax</i>	hoax	<i>hoax</i>
moat	<i>moat</i>	moat	<i>moat</i>
peach	<i>peach</i>	peach	<i>peach</i>
snack	<i>snack</i>	snack	<i>snack</i>
vogue	<i>vogue</i>	vogue	<i>vogue</i>

Fig. 1 Examples of stimuli from each script condition.

instrument resembles a normal pen, with the addition of a small camera protruding from the tip. The camera reads a fine dot pattern on the paper, generating a digital trace of each pen stroke. The stimulus images were enlarged (to match the size of the computer-generated words) and sharpened using Adobe Photoshop. Mulatti et al. (2006) erroneously included one stimulus (“girl”) twice in the dense word list. We excluded this duplicate, leaving 30 words in the sparse neighborhood condition and 29 in the dense neighborhood condition. The stimuli were originally balanced on 11 additional variables: phonological neighborhood frequency, orthographic neighborhood density, orthographic neighborhood frequency, word frequency, feedforward consistency, feedback consistency, letter length, regularity, whammies, imageability, and age of acquisition (see Mulatti et al. 2006, for the complete details).

Apparatus Stimuli were presented via the E-Prime 1.2 program (Schneider, Eschman, & Zuccolotto, 2002) on a Dell computer with a plasma-screen monitor. Naming response times (RTs) were collected by a standard voice key connected to an E-Prime SR response box.

Procedure The experiment began with a brief naming task wherein participants read the numbers 1 through 10 aloud. This warm-up task acclimated participants to the sensitivity of the voice key. The procedure continued with nine practice word-naming trials, then 59 experimental trials. Each trial began with a fixation point (***) appearing at the center of the screen for 750 ms. The fixation point was replaced by the target word, which remained until a vocal response was initiated. The next trial began after a 1,000-ms interval. If no response was detected within 3 s, the experiment moved on to the intertrial interval. An experimenter recorded any mispronunciations and voice key errors as they occurred. The 59

words were presented in a random order for each participant, all in the same script.

Results and discussion

Correct RTs Seven participants were excluded from the analyses (three from the computer print condition, one from the computer cursive condition, one from the natural print condition, and two from the natural cursive condition) due to having either average RTs that were more than three standard deviations above the respective group means or error rates more than three standard deviations greater than average. Trials with mispronunciations or voice key errors were removed from the RT data prior to the analysis, constituting 3.8% and 2.6% of trials, respectively. RTs greater than three standard deviations from the group means were replaced with cutoff scores (following Winer, 1971), resulting in the replacement of 1.4% of the correct RTs. Mean RTs and the derived neighborhood effects per condition are shown in Table 1.

Across all analyses, we employed linear mixed models, including subjects and items as random effects (Locker, Hoffman, & Bovaird, 2007). We first examined each script condition individually with the factor Phonological Neighborhood Density (dense, sparse). The computer print condition produced a marginal 12-ms neighborhood effect, $F(1, 56.81) = 3.54, p = .065, r^2_{\text{pseudo}} < .01$. The neighborhood effects did not approach significance in any of the other script conditions.

We followed the separate analyses with an omnibus model including the factors Script (computer print, computer cursive, natural print, natural cursive) and Neighborhood (dense, sparse). A significant script effect was apparent, $F(3, 242.31) = 17.34, p < .001, r^2_{\text{pseudo}} = .16$, with RTs increasing with degrees of physical ambiguity. The neighborhood effect was null, but a reliable Script $\times$ Neighborhood interaction did emerge, $F(3, 8291.52) = 3.22, p = .02, r^2_{\text{pseudo}} < .01$, reflecting a general reduction in the neighborhood effect with the addition of physical ambiguity.

To further examine how natural ambiguity influenced neighborhood effects, we contrasted the computer and

natural conditions in a model with the factors Source (computer, human) and Neighborhood. We observed a reliable source effect, $F(1, 277.69) = 18.92, p < .001, r^2_{\text{pseudo}} = .01$, with higher RTs for the naturally produced stimuli. We also found a Source $\times$ Neighborhood interaction, with a larger neighborhood effect for computer-generated words, $F(1, 8299.28) = 6.17, p = .01, r^2_{\text{pseudo}} < .01$.

Error rates Error rates (see Table 1) were generally low across all conditions. The omnibus model produced only a reliable script effect, $F(3, 169.76) = 14.19, p < .001, r^2_{\text{pseudo}} = .11$; errors increased with greater physical ambiguity.

Experiment 2: Orthographic neighborhood effects

Experiment 2 was a replication and extension of Experiment 1 from Mulatti et al. (2006), examining orthographic neighborhood effects using the same four scripts as in Experiment 1. We again expected the computer-generated print condition to replicate the findings from Mulatti et al. (2006), with a negligible neighborhood effect. We also expected this effect to become more robust when human-generated scripts were used, since top-down lexical activation could better disambiguate features at the letter level.

Method

Subjects The observers were 174 Arizona State University students who received course credit. Of these volunteers, 41 were in the computer print, 46 in the computer cursive, 44 in the natural print, and 43 in the natural cursive condition.

Stimuli The stimuli from Experiment 2 of Mulatti et al. (2006) were generated in the formats used in Experiment 1. Sixty words were used, half from dense orthographic neighborhoods and half from sparse orthographic neighborhoods. The stimuli were balanced on 11 additional variables, as in Experiment 1 (see Mulatti et al., 2006).

Table 1 Experiment 1: Mean response times (and error rates in parentheses) by script and phonological neighborhood density

Script	Phonological neighborhood density		Neighborhood effect (ms)
	Dense	Sparse	
Computer print	524 (.03)	536 (.03)	12
Computer cursive	618 (.08)	641 (.07)	22
Natural print	572 (.05)	580 (.06)	8
Natural cursive	688 (.12)	680 (.11)	−9

Mean neighborhood effects (sparse minus dense) are shown in the third column

Apparatus and procedure The apparatus and procedure were identical to those of Experiment 1.

Results and discussion

Correct RTs Five participants were excluded from the analyses (one from the computer print, two from the computer cursive, one from the natural print, and one from the natural cursive condition) due to having either average RTs or error rates more than three standard deviations above the respective group mean. One word (“joist”) was excluded from all conditions for eliciting extremely high error rates. Trials with mispronunciations or voice key errors were removed from the RT data prior to the analysis, constituting 3.5% and 1.9% of trials, respectively. The data were trimmed in the same manner as in Experiment 1, resulting in the replacement of 1.3% of the correct RT data. Mean RTs and the derived neighborhood effects per condition are shown in Table 2.

We first examined each script condition individually via linear mixed models, with Subjects and Items as random factors and the fixed factor Orthographic Neighborhood Density (dense, sparse). The computer print condition replicated the findings of Mulatti et al. (2006), producing a nonsignificant 10-ms neighborhood effect. The neighborhood effect was not reliable in the computer cursive condition, but the natural print condition produced a significant 25-ms effect, $F(1, 55.27) = 4.54, p = .04, r^2_{\text{pseudo}} < .01$. Finally, the natural cursive condition produced a 19-ms neighborhood effect that was not reliable.

We created an omnibus model, including the factors Script (computer print, computer cursive, natural print, natural cursive) and Orthographic Neighborhood (dense, sparse). A reliable script effect emerged, $F(3, 254.11) = 10.55, p < .001, r^2_{\text{pseudo}} = .11$, with slower RTs for naturally produced words. The neighborhood effect was not reliable, but we did observe a Script $\times$ Neighborhood interaction, $F(3, 9169.20) = 3.19, p = .02, r^2_{\text{pseudo}} < .01$: Generally, the neighborhood effects were larger for words presented in naturally produced scripts (see Table 2).

We next compared the computer-generated and natural scripts in a model with the factors Source (computer, human) and Neighborhood. There was a reliable effect of source, $F(1, 270.55) = 18.63, p < .001, r^2_{\text{pseudo}} = .04$, with slower RTs for naturally produced words. A null effect of neighborhood was qualified by a significant Source $\times$ Neighborhood interaction, $F(1, 8889.78) = 7.18, p = .007, r^2_{\text{pseudo}} < .01$. Critically, the neighborhood effect was larger for words in naturally produced scripts.

Finally, we combined the data from Experiments 1 and 2 in a large model with the factors Experiment (1, 2), Script (computer print, computer cursive, natural print, natural cursive), and Neighborhood (dense, sparse) to assess the relative effect that each script had upon orthographic and phonological

neighborhood effects. This analysis produced a marginal experiment effect, $F(1, 16921.48) = 3.46, p = .02, r^2_{\text{pseudo}} < .01$, with faster RTs in Experiment 1. We also found a script effect, $F(3, 391.79) = 19.71, p < .001, r^2_{\text{pseudo}} = .13$, with higher RTs for naturally produced scripts. A neighborhood effect, $F(1, 62.13) = 21.10, p < .001, r^2_{\text{pseudo}} < .01$, reflected faster responding for words from dense neighborhoods. These main effects were qualified by a couple of interactions. First, an Experiment $\times$ Script interaction, $F(3, 16833.98) = 7.19, p < .001, r^2_{\text{pseudo}} < .01$, was driven by increased RTs for the natural print condition in Experiment 2, relative to Experiment 1. Importantly, a three-way Experiment $\times$ Script $\times$ Neighborhood interaction also emerged, $F(3, 17581.16) = 5.60, p = .001, r^2_{\text{pseudo}} < .01$: Phonological neighborhood effects decreased when moving from artificial to natural stimuli, whereas orthographic neighborhood effects increased.¹

Error rates Again, error rates were generally low across all conditions. The omnibus analysis with the factors Script and Neighborhood produced a marginal main effect of script, $F(3, 180.22) = 2.56, p = .057, r^2_{\text{pseudo}} < .01$, with accuracy decreasing for natural stimuli. There was also a Script $\times$ Neighborhood interaction, $F(3, 9742.23) = 4.11, p = .006, r^2_{\text{pseudo}} < .01$, due to lower accuracy in the natural cursive condition for words from sparse neighborhoods.

General discussion

As we anticipated, both of the present experiments replicated the findings from Mulatti et al. (2006) when computer-generated text was used, but reversed their findings when natural, handwritten text was used. Given computer text in Experiment 1, a facilitatory phonological neighborhood effect was observed, but it disappeared with the addition of physical ambiguity in handwriting. In Experiment 2, this pattern reversed: With computer-generated words, orthographic neighborhood effects were negligible, but these grew robust for handwritten words. The nonsignificant 2-ms neighborhood effect observed by Mulatti et al. (2006) with typewritten words was comparable to the null 10-ms effect that we observed. Although these very weak trends suggest that orthographic neighborhoods exert little influence in word perception, we contend that, given pristine text as input, letter-level processing is too “close to ceiling” for lexical feedback to have clear effects. When we presented the same words in

¹ The same analysis was carried out on z-transformed data within subjects, such that the data across experiments were placed on the same scale. The three-way Experiment $\times$ Script $\times$ Neighborhood interaction was still reliable in this more conservative analysis.

Table 2 Experiment 2: Mean response times (and error rates in parentheses) by script and orthographic neighborhood density

Script	Orthographic neighborhood density		Neighborhood effect (ms)
	Dense	Sparse	
Computer print	558 (.05)	569 (.05)	10
Computer cursive	612 (.05)	613 (.04)	2
Natural print	586 (.05)	611 (.05)	25*
Natural cursive	664 (.05)	684 (.08)	19

Mean neighborhood effects (sparse minus dense) are shown in the third column. * $p < .05$

handwritten form, the neighborhood effects became far stronger.

The use of pristine text across experiments confers many benefits, including the control of extraneous variation and generality across laboratories. Nonetheless, as has been noted previously (Barnhart & Goldinger, 2010; Manso de Zuniga et al., 1991), the entire empirical database in printed word perception was derived from experiments using ambiguity-free text. This differs starkly from spoken word perception, wherein higher-level context is often necessary to allow for recovery of segmental units (e.g., Metsala & Walley, 1998). Even if word perception is truly a highly interactive process (e.g., Plaut, McClelland, Seidenberg, & Patterson, 1996; Van Orden & Goldinger, 1994), we might systematically underestimate the role of feedback processes by routinely giving them too little work to perform.

In summarizing their results, Mulatti et al. (2006) called for the revision of three well-known models of word recognition, which were relatively incapable of simulating their observed phonological neighborhood effects or the absence of orthographic neighborhood effects.² In the present research, handwritten words (in both printed and cursive styles) produced a reversal of these findings, creating an empirical profile that better fits the original models. Taken together, the results suggest that, in any model of word perception, the onus is placed on whichever processes must deal most directly with ambiguity. Given pristine visual input, the greatest ambiguity arises in matching orthography to phonology, leading to phonological neighborhood effects. Given inconsistent, naturally produced items, the greatest ambiguity occurs in matching features to letters, leading to orthographic neighborhood effects.

The phonological neighborhood effect observed by Mulatti et al. (2006) is not particularly surprising, considering other crossmodal interactions that have been reported. Ziegler and Muneaux (2007) found that orthographic neighborhoods

affect *spoken* word perception when people are proficient readers (excluding, e.g., children and dyslexics). In contrast, phonological neighborhood effects were apparent across levels of reading proficiency. Ziegler and Muneaux suggested that a restructuring of phonological word representations occurs with literacy development, as has been suggested by *lexical restructuring* theory (Metsala & Walley, 1998), with orthographic knowledge allowing more fine-grained coding of words from dense phonological neighborhoods (which would be easily confused in the auditory modality). The development of this fine-grained coding increases feedback effects on lower-level perception, both from sound to print and from print to sound. This crossmodal feedback may underlie some curious findings in word perception, such as the feedback consistency effect (Stone, Vanhoy, & Van Orden, 1997), wherein lexical decision is slowed by seemingly irrelevant alternative spellings activated by derived phonology. Similar feedback processes might underlie the phonological neighborhood effect observed with typewritten words in Experiment 1 (and by Mulatti et al., 2006). Given unambiguous input, it becomes possible for smaller feedback effects from phonology to affect naming times. With handwritten words, the processing required to link pen strokes to the orthographic lexicon swamps this relatively weak feedback from alternative phonologies.

On the surface, the present results represent an empirical counterpoint to the previous observation by Mulatti et al. (2006). But they also help emphasize a more important theoretical question: What exactly should we hope to achieve when creating models of word recognition? Perhaps models should be continually refined to account for every small effect that is reliably observed under ideal, laboratory conditions. The alternative (and possibly more compelling) answer is that models should reflect *natural* human performance. On a daily basis, people encounter written information in many forms, and a great deal is likely analogous to that in the laboratory (e.g., reading news stories online). Nevertheless, people are capable of far more complex perceptual feats, reading words that are neither pristine nor unambiguous: We read signs at odd angles and

² In a rebuttal to Mulatti et al. (2006), Perry, Ziegler, and Zorzi (2007) demonstrated that the CDP+ model could replicate the effects (and null effects) reported in the earlier study, using the same stimuli.

orientations, journal articles that are duplicated with poor quality, and handwritten notes of all kinds. The perceptual system routinely deals with degraded visual input. Exploring perception with more natural stimuli might lead to deeper theoretical insights.

Author note Support was provided by NIH Grant Number R01 HD075800-02 to S.D.G. We thank Megan Papesh and Melissa Gillen for generating the stimulus materials.

References

- Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 802–814. doi:10.1037/0278-7393.15.5.802
- Barnhart, A. S., & Goldinger, S. D. (2010). Interpreting chicken-scratch: Lexical access for handwritten words. *Journal of Experimental Psychology: Human Perception and Performance*, 36, 906–923. doi:10.1037/a0019258
- Becker, C. A., & Killion, T. H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 389–401. doi:10.1037/0096-1523.3.3.389
- Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI: Proceedings of the Sixth International Symposium on Attention and Performance* (pp. 535–555). Hillsdale: Erlbaum.
- Huntsman, L. A., & Lima, S. D. (2002). Orthographic neighbors and visual word recognition. *Journal of Psycholinguistic Research*, 31, 289–306.
- Locker, L., Jr., Hoffman, L., & Bovaird, J. A. (2007). On the use of multilevel modeling as an alternative to items analysis in psycholinguistic research. *Behavior Research Methods*, 39, 723–730. doi:10.3758/BF03192962
- Manso De Zuniga, C., Humphreys, G. W., & Evett, L. J. (1991). Additive and interactive effects of repetition, degradation, and word frequency in the reading of handwriting. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 10–33). Hillsdale: Erlbaum.
- McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, 86, 287–330. doi:10.1037/0033-295X.86.4.287
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, 88, 375–407. doi:10.1037/0033-295X.88.5.375
- Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Mahwah: Erlbaum.
- Mulatti, C., Besner, D., & Job, R. (2003, November). *Reading aloud: Evidence for a phonologically based neighborhood effect*. Paper presented at the 44th Annual Meeting of the Psychonomic Society, Vancouver, BC, Canada.
- Mulatti, C., Reynolds, M. G., & Besner, D. (2006). Neighborhood effects in reading aloud: New findings and new challenges for computational models. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 799–810. doi:10.1037/0096-1523.32.4.799
- Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review*, 114, 273–315. doi:10.1037/0033-295X.114.2.273
- Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103, 56–115. doi:10.1037/0033-295X.103.1.56
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh: Psychology Software Tools, Inc.
- Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 876–900. doi:10.1037/0096-1523.21.4.876
- Stone, G. C., Vanhoy, M., & Van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language*, 36, 337–359.
- Van Orden, G. C., & Goldinger, S. D. (1994). Interdependence of form and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 1269–1291. doi:10.1037/0096-1523.20.6.1269
- Winer, B. J. (1971). *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill.
- Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's *N*: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, 15, 971–979. doi:10.3758/PBR.15.5.971
- Ziegler, J. C., & Muneaux, M. (2007). Orthographic facilitation and phonological inhibition in spoken word recognition: A developmental study. *Psychonomic Bulletin & Review*, 14, 75–80. doi:10.3758/BF03194031