Psychon Bull Rev (2011) 18:605–611
DOI 10.3758/s13423-011-0078-9

Observing neighborhood effects without neighbors

Lidia Suárez · Seok Hui Tan · Melvin J. Yap ·
Winston D. Goh

Abstract With a new metric called *phonological Levenshtein distance* (PLD20), the present study explores the effects of phonological similarity and word frequency on spoken word recognition, using polysyllabic words that have neither phonological nor orthographic neighbors, as defined by neighborhood density (the *N*-metric). Inhibitory effects of PLD20 were observed for these lexical hermits: Close-PLD20 words were recognized more slowly than distant PLD20 words, indicating lexical competition. Importantly, these inhibitory effects were found only for low- (not high-) frequency words, in line with previous findings that phonetically related primes inhibit recognition of low-frequency words. These results indicate that the properties of PLD20—a continuous measure of word-form similarity—make it a promising new metric for quantifying phonological distinctiveness in spoken word recognition research.

Keywords Spoken word recognition · Word recognition · Psycholinguistics

It is well established in the literature that similarity between word forms is influential in determining the speed and accuracy with which speech stimuli are identified. One way in which research on spoken word recognition has characterized word-form similarity is phonological neighborhood density (*N*-metric: Luce & Pisoni, 1998). Neighborhood density—the number of words that differ by a phoneme from the target word—is higher for a word like *cat*, which has many neighbors (e.g., *hat*, *cut*, *at*, *catty*), than for a word like *wag*, which has fewer neighbors (e.g., *bag*, *wan*). Using auditory lexical decision, wherein participants indicate whether stimuli are words or nonwords, previous studies have shown that words from dense neighborhoods are recognized more slowly than words from sparse ones (e.g., Goh, Suárez, Yap, & Tan, 2009; Luce & Pisoni, 1998; Ziegler, Muneaux, & Grainger, 2003).

Similar effects of neighborhood density have also been found with other variants of word-form similarity measures based on the single-phoneme-change metric, such as the clustering coefficient (*C*-metric: Watts & Strogatz, 1998) and neighborhood spread (*P*-metric: Andrews, 1997). Words with many neighbors that are also neighbors to one another (i.e., words with high coefficients) are recognized more slowly than words with the same number of neighbors, of which few are neighbors to one another (Chan & Vitevitch, 2009). Similarly, words with more phonemes that can be changed to form neighbors (i.e., high spread) are recognized more slowly than words with fewer such phonemes (Vitevitch, 2007). Along with evidence that words with many neighbors sharing common onsets are recognized more slowly than words with fewer such neighbors (Magnuson, Dixon, Tanenhaus, & Aslin, 2007), these findings indicate that word-form similarity gives rise to lexical competition, which occurs during word identification.

By database estimates, many English words have no neighbors, as estimated using traditional definitions of density. In the English Lexicon Project (ELP; Balota, Yap, Cortese, Hutchison, Kessler, Loftis, et al., 2007), these words include 2.5% of monosyllabic, 43.9% of disyllabic, and 76.3% of trisyllabic words. Despite many words having no neighbors, the evidence for current models of spoken word recognition (e.g., Luce & Pisoni, 1998) has so far been based on monosyllabic and disyllabic words that do

L. Suárez (✉) · S. H. Tan · M. J. Yap · W. D. Goh
Department of Psychology, Faculty of Arts and Social Sciences,
National University of Singapore,
Block AS4, 9 Arts Link,
Singapore 117570, Singapore
e-mail: suarez.l@nus.edu.sg

have neighbors (e.g., Vitevitch & Luce, 1999; Vitevitch, Stamer, & Sereno, 2008). Indeed, lexical hermits, which are words with no neighbors, represent the largest part of the lexicon in Vitevitch's (2008) analyses. Theoretical accounts of word recognition should therefore not be based solely on phenomena observed in a restricted set of words in the lexicon (words with traditional neighbors). It is plausible that hermits have differential levels of word-form similarity, and are thus also subject to different grades of lexical competition. However, the current *N*-, *P*-, and *C*-density metrics cannot capture these effects, because they treat hermits identically. These metrics make no provision to distinguish hermits that share more word-form similarity with other words in the lexicon from hermits that share less similarity. Both these types of hermits are treated as isolated by the failure to find any neighbor based on a single phoneme difference.

In relying on density metrics, previous studies have not examined the effects of word-form similarity on word recognition where there are no neighbors. Despite this, one earlier study has addressed this issue by indirectly priming monosyllabic targets with words that share phonetic features (but not phonemes) with the targets (Goldinger, Luce, & Pisoni, 1989). By using estimates of phonetic similarity derived from confusion matrices of consonants and vowels in noise, the authors demonstrated that related primes (words phonetically related but not neighbors to targets) inhibited target recognition accuracy relative to unrelated primes (words that were neither phonetically related nor neighbors to targets). The inhibitory priming effects observed were attributed to the transient activation of the target word's neighborhood by the related "non-neighbor" prime (see also Goldinger, Luce, Pisoni, & Marcario, 1992; Luce, Goldinger, Auer, & Vitevitch, 2000). However, this approach assumes that primes are needed to indirectly boost the neighborhood of target words in order to observe word similarity effects.

Word-form similarity effects can be examined without primes by using a different metric. This new metric of word-form similarity, known as *phonological Levenshtein distance* or *PLD20* (Yap & Balota, 2009; Yarkoni, Balota, & Yap, 2008), makes possible a study of distant neighborhood effects in lexical hermits. PLD20 reflects the mean number of steps required through phoneme substitutions, insertions, or deletions (Levenshtein distance) to transform a word into its 20 closest Levenshtein neighbors in the ELP lexicon (Balota et al., 2007). PLD20 values are smaller for close-PLD20 words, such as *resume*, than for distant-PLD20 words, such as *insomnia*. As illustrated in Fig. 1, *resume*, which transforms into its closest Levenshtein neighbors (e.g., *result*, *refute*) in approximately two steps on average, is more similar to its Levenshtein neighbors than is *insomnia*, which transforms into its closest Levenshtein neighbors (e.g., *inertia*, *insolent*) in approximately four steps on average. Levenshtein distance, which is synonymous with edit distance (see Bailey & Hahn, 2001), is a continuous measure of lexical similarity. As with density, in which words from dense neighborhoods are recognized more slowly than words from sparse ones, close-PLD20 words should be recognized more slowly than distant-PLD20 words. Unlike density, however, PLD20 can be applied to all words, particularly longer words that may not have traditional neighbors.

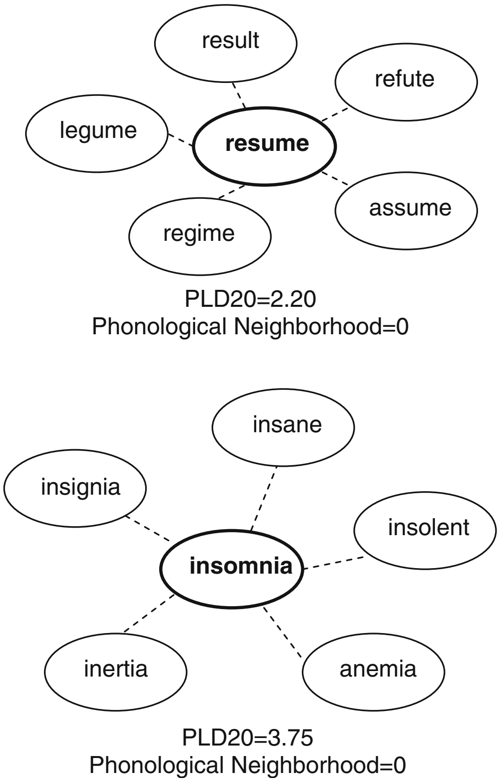

Fig. 1 Examples of PLD20 neighborhoods. The top panel depicts some PLD20 neighbors of a close-PLD20 word (in bold), and the bottom panel some PLD20 neighbors of a distant PLD20 word. The length of the links between the word and its neighbors represents the number of changes by deletion, insertion, or substitution that are necessary to transform the target into the neighbor, so longer links represent more changes. The figure depicts only some neighbors, but the total PLD20 value is the average of the 20 closest PLD20 neighbors. Note that words have PLD20 neighbors even when their phonological neighborhood density is zero as measured by the traditional method of a single phoneme difference

Although previous studies on visual word recognition have shown that words with closer PLD20 neighbors are recognized more slowly than words with distant PLD20 neighbors (Yap & Balota, 2009), and that PLD20 is superior to density for predicting response latencies for longer words (Yarkoni et al., 2008), no previous study has examined the effects of PLD20 on spoken word recognition. More importantly, the effects of word-form similarity on word recognition where words have no neighbors have yet to be explored. To address this issue, we examined the effects of PLD20 and word frequency on response latencies (RTs) using an auditory lexical decision task (LDT). We predicted main effects of PLD20 and word frequency: Recognition should be slower for close- than for distant-PLD20 words, and for low- than for high-frequency words. Density effects have been observed to be stronger for high- than for low-frequency monosyllabic words (Goh et al., 2009; Luce & Pisoni, 1998), suggesting that PLD20 effects should be stronger for high- than for low-frequency words. However, Yarkoni et al. (2008) found orthographic Levenshtein distance (OLD20) effects to be stronger for low- than for high-frequency words in visual word recognition. This has been attributed to the weaker representation of low-frequency words, which makes them more susceptible to lexical competition. Furthermore, Goldinger et al. (1989) observed that for low-density targets (similar to words with no traditional neighbors), priming with low-frequency, phonetically related words (but not with their high-frequency counterparts) inhibited target recognition. These findings suggest that a similar overadditive interaction pattern between word frequency and word-form similarity might be observed in the present study. Determining whether the interaction between PLD20 and word frequency is overadditive or underadditive is thus another objective of this study.

Method

Participants

Seventy-two[1] students from the National University of Singapore with no known speech or hearing impairments participated for course credit.

Design and materials

Using a 2 (PLD20: close, distant) x 2 (frequency: low, high) within-subjects design, 38 disyllabic and trisyllabic words with no orthographic or phonological neighbors (using the ELP values: Balota et al., 2007) and an equivalent number of nonwords were presented in each condition (see the Appendix for the stimuli). Nonwords were constructed by replacing the last one or two phonemes of real words (not used as stimuli) to ensure that participants listened to the entire tokens before responding (Vitevitch, 2008). Table 1 shows that the words differed on PLD20, $F(1, 148) = 168.40$, $MSE = 0.141$, $p < 0.001$ (close PLD20, $M = 2.42$, $SD = 0.27$; distant PLD20, $M = 3.21$, $SD = 0.46$), and log frequency (using the CELEX database), $F(1, 148) = 181.70$, $MSE = 0.12$, $p < 0.001$ (low frequency, $M = 0.08$, $SD = 0.12$; high frequency, $M = 1.05$, $SD = 0.62$). Table 2 shows that words were equated for uniqueness point (Luce, 1986), phoneme and syllable length, and duration[2] (all *F*s < 1.88).

Table 1 Mean log frequency and PLD20 of words across conditions

	Log Frequency		PLD20	
Conditions	*M*	*SD*	*M*	*SD*
Low Frequency				
Close PLD20	0.09	0.12	2.49	0.25
Distant PLD20	0.06	0.11	3.19	0.51
High Frequency				
Close PLD20	1.06	0.60	2.35	0.27
Distant PLD20	1.04	0.63	3.23	0.41

PLD20 scores and log frequency were obtained from the ELP (Balota et al., 2007) and CELEX databases, respectively

The stimuli were recorded by a female Singaporean using 16-bit mono, 44.1-kHz, .wav-format recording, with overall root-mean-square amplitudes digitally leveled. The mean correct-identification levels for words and nonwords by 27 undergraduates from the same population sample (who did not participate in the study) were 95% (*SD* = 7%) and 86% (*SD* = 12%), respectively.

Procedure

Participants were tested on individual PCs in groups of 5 or fewer, using E-Prime 1.2 and PST Serial Response Boxes (Schneider, Eschman, & Zuccolotto, 2002) with the left- and rightmost buttons labeled *nonword* and *word*, respectively. Stimuli were binaurally played through Beyerdynamic DT150 headphones at approximately 70 dB SPL. Participants were asked to indicate as quickly and as accurately as possible whether or not each stimulus was a real word. RTs were measured from stimulus onset to buttonpress. Twenty practice trials were presented using stimuli unrelated to the study, followed by 304 experimental trials, with a 500-ms intertrial interval and in a random order for each participant, with a short break after half of the trials were completed.

---

[1] Two identical experiments, initially conducted with 45 participants in the main experiment and 22 in the replication (after eliminating outliers), yielded virtually identical results. Hence, the data sets were collapsed and presented as a single study.

[2] Although duration was equated on the basis of the nonsignificant interaction in the omnibus ANOVA [$F(1, 148) = 1.53$, $MSE = 10{,}326.64$, $p = 0.22$], we subsequently noticed a 46-ms difference between close-PLD20 low-frequency and distant-PLD20 low-frequency words that was significant when contrasted [$F(1, 74) = 4$, $MSE = 9{,}977.68$, $p = 0.049$]. This is addressed statistically in the Results section.

Table 2 Psycholinguistic characteristics of words across conditions

Conditions	Uniqueness Point	No. of Phonemes	No. of Syllables	Word Duration (ms)
Low Frequency				
Close PLD20	5.61 (1.35)	6.34 (0.85)	2.55 (0.50)	781 (103)
Distant PLD20	5.21 (1.34)	6.53 (0.86)	2.45 (0.50)	827 (97)
High Frequency				
Close PLD20	5.58 (1.13)	6.50 (0.69)	2.58 (0.50)	809 (108)
Distant PLD20	5.24 (1.34)	6.76 (0.71)	2.50 (0.51)	814 (98)
Nonwords		6.82 (1.53)	2.51 (0.50)	812 (119)

*SD*s are in parentheses

Results

We computed overall word and nonword means and *SD*s for each participant. Data for trials on which the response latency exceeded 2.5 *SD*s from the participant's individual mean were replaced with a 2.5-*SD* cutoff value.[3] The data for 7 participants who obtained response accuracy means 2.5 *SD*s below the overall accuracy mean were discarded in subsequent analyses. Thus, analyses were carried out with data from 67 participants.

The latencies by participants and by items are summarized in Table 3. Raw RTs and ANCOVA results using the raw RTs are reported in the two left columns;[4] RTs adjusted according to Luce and Pisoni's (1998) method of subtracting word-token duration from the RT for that word are reported in the two right columns. All two-way ANOVAs reported below are based on the adjusted RTs.

For latency, the PLD20 main effect was significant by participants, $F_p(1, 66) = 16.76$, $MSE = 1{,}489.68$, $p < 0.001$, but not by items, $F_i(1, 148) = 1.68$, $MSE = 8{,}888.35$, $p = 0.197$; the frequency main effect was reliable, $F_p(1, 66) = 156.68$, $MSE = 1{,}268.59$; $F_i(1, 148) = 14.91$, $MSE = 8{,}888.35$, $ps < 0.001$. The main effects were qualified by an interaction that was significant by participants, $F_p(1, 66) = 48.46$, $MSE = 1{,}028.68$, $p < 0.001$, and marginally significant by items, $F_i(1, 148) = 3.64$, $MSE = 7{,}987.51$, $p = 0.057$. Simple main effects showed that latencies were shorter for high- than for low-frequency words at both the close-PLD20 level, $F(1, 66) = 147.47$, $MSE = 1{,}517.86$, and the distant-PLD20 level, $F(1, 66) = 31.78$, $MSE = 779.41$, $ps < 0.001$. The interaction was driven by the finding that with low-frequency words, participants were faster for distant- than for close-PLD20 words, $F(1, 66) = 43.99$, $MSE = 1{,}652.49$, $p < 0.001$; in contrast, there was a null PLD20 effect for high-frequency words, $F(1, 66) = 2.46$, $MSE = 865.87$, $p = 0.122$.

Accuracy is summarized in Table 4. There was a PLD20 main effect by participants, $F_p(1, 66) = 7.24$, $MSE = 0.002$, $p = 0.009$, but not by items, $F_i < 1$, and a main effect of frequency, $F_p(1, 66) = 118.83$, $MSE = 0$, $p < 0.001$; $F_i(1, 148) = 21.13$, $MSE = 0.012$, $p < 0.001$. The effects were qualified by a significant interaction, $F_p(1, 66) = 37.10$, $MSE = 0.002$, $p < 0.001$; $F_i(1, 148) = 4.41$, $MSE = 0.012$, $p = 0.037$. Simple effects showed that accuracy was higher for high- than for low-frequency words at the close-PLD20 level by both participants and items, $F_p(1, 66) = 126.75$, $MSE = 0$; $F_i(1, 148) = 22.42$, $MSE = 0.012$, $ps < 0.001$, and at the distant-PLD20 level by participants, $F_p(1, 66) = 21.88$, $MSE = 0$, $p < 0.001$, but not by items, $F_i < 1$. For low-frequency words, accuracy was higher for distant- than for close-PLD20 words, $F_p(1, 66) = 25.58$, $MSE = 0$, $p < 0.001$; $F_i(1, 148) = 4.07$, $MSE = 0.012$, $p = 0.045$. However, for high-frequency words, the reverse was true only by participants, $F_p(1, 66) = 17.25$, $MSE = 0$, $p < 0.001$, but not by items, $F_i < 1$. As shown in Table 3, the interaction effect, which is likely related to the small *SE*s found for some conditions, is minimal (−2), and therefore is not discussed further.

Discussion

The noteworthy finding is that PLD20 predicts latencies in spoken word recognition for polysyllabic words, even when these words have neither phonological nor orthographic neighbors as defined by the *N*-metric. In addition to the robust finding that recognition is slower for low- than for high-frequency words, our results demonstrate that recognition is slower for close- than for distant-PLD20 words. Consistent with current models of spoken word recognition, such as the

---

[3] Analyses performed with no trimming procedures, or with the elimination of trials in which RTs exceeded 2.5 *SD*s of participants' means, yielded similar results.

[4] The ANOVA by participants conducted using raw RTs showed that distant-PLD20 words were recognized more slowly than close-PLD20 words [$F_p(1, 66) = 3.12$, $MSE = 1{,}515.54$, $p = 0.08$] and revealed no interaction [$F_p(1, 66) = 1.91$, $MSE = 954.68$, $p = 0.17$]. Although an ANCOVA by items, controlling for word duration using the raw RTs, revealed no PLD20 main effect ($F_i < 1$) and no interaction, $F_i(1, 147) = 2.18$, $MSE = 6{,}195.77$, $p = 0.14$, the trends were in the same direction as those found with the adjusted RTs. This suggests that the discrepancies initially observed in the raw RTs can be attributed to word duration differences between close- and distant-PLD20 words in the low-frequency condition. As shown in Table 3, the PLD20 effect in the low-frequency condition reverses in the by-participants and ANCOVA analyses. We thus reported results using the adjusted RTs, so that analyses by participants and by items could be made on the same dependent variable. The same substituting procedure using 2.5-*SD* cutoff values was employed with the adjusted RTs.

Table 3 Mean latencies across word frequency and PLD20

Conditions	Raw RTs (participants)	ANCOVA (items)	Adjusted RTs (participants)	Adjusted RTs (items)
Low Frequency				
Close PLD20	992.38 (10.24)	1017.44 (12.88)	214.52 (10.34)	223.59 (14.72)
Distant PLD20	995.55 (10.67)	991.95 (12.83)	167.94 (10.62)	174.46 (15.09)
PLD20 effect	–3.17	25.49	46.58	49.13
High Frequency				
Close PLD20	941.06 (9.09)	943.42 (12.77)	132.77 (9.03)	135.23 (15.36)
Distant PLD20	954.67 (9.91)	955.84 (12.78)	140.74 (9.82)	144.72 (15.98)
PLD20 effect	–13.61	–12.42	–7.97	–9.49
Interaction	10.44	13.07	38.61	39.64
Nonwords	1,110.15 (17.45)	1,111.12 (7.27)	300.33 (17.44)	300.38 (7.48)

*SE*s are in parentheses

neighborhood activation model, that predict that word-form similarity induces lexical competition during spoken word recognition (Luce & Pisoni, 1998), the results indicate that PLD20 is a valid measure of phonological similarity.

With respect to the interaction between PLD20 and word frequency, we found that PLD20 effects were larger for low- (47 ms) than for high-frequency words (8 ms). This is similar to the pattern reported by Yarkoni et al. (2008), where larger effects of orthographic Levenshtein distance (OLD20) were observed for low- than for high-frequency words in visual word recognition. The underlying assumption is that Levenshtein-based neighborhood measures reflect global word-form similarity, whereby words with closer neighbors are more likely than words with distant neighbors to be pulled into attractor basins containing orthographically or phonologically similar words (depending on whether OLD20 or PLD20 is being used). It can be argued that PLD20 effects are stronger for low- than for high-frequency words because the weak representations of low-frequency words make them more easily influenced by the attraction of the phonological basin, as compared to high-frequency words, which have relatively stronger representations.

The distant neighborhood effects observed in the present study are also consistent with Goldinger et al.'s (1989) findings that accuracy in auditory perceptual identification was compromised more by related primes (which are phonetically similar but not neighbors to targets) than by unrelated primes (which are neither phonetically similar nor neighbors to targets). Together, the results show that distant neighbors—neither PLD20 neighbors nor related primes are traditional neighbors to target words—hinder word recognition through lexical competition. Intriguingly, the observation that PLD20 effects were larger for low- than for high-frequency words in our study meshes well with previous findings. Goldinger et al. observed that when targets were from sparse neighborhoods (much like the lexical hermits in our study), the inhibitory effects of prime frequency (poorer accuracy for low- than for high-frequency primes) were much stronger for low- than for high-frequency targets. Both findings reflect greater activation and competition from phonetically similar candidates. The similarities between our close-PLD20 condition and their low-frequency prime condition allow us to draw parallels between the larger effects of PLD20 for low-frequency words observed here and the larger prime frequency effects on low-frequency targets obtained by Goldinger and colleagues. The present findings may serve to clarify the previous results: Distant neighborhood effects can be observed without the additional activation of distant neighbors through related primes, but from the mere presence of these distant (PLD20) neighbors.

Table 4 Mean accuracy across word frequency and PLD20

	Accuracy	
Conditions	*M*	*SE*
Low Frequency		
Close PLD20	86	1
Distant PLD20	91	1
PLD20 effect	–5	
High Frequency		
Close PLD20	97	0
Distant PLD20	94	1
PLD20 effect	3	
Interaction	–2	
Nonwords	85	1

We note that the interaction found does not, on the surface, appear consistent with the interaction between neighborhood density and word frequency observed in previous studies. Specifically, Goh et al. (2009) found larger density effects for high- (68 ms) than for low- (30 ms) frequency words, a pattern also reported by Luce and Pisoni (1998), who found effects of 21 versus 6 ms, respectively. If one assumes that PLD20 effects are analogous to density effects, why are density effects stronger for high-frequency words? Before we address this question, we need to point out that the dense and sparse neighborhood conditions in Goh et al.'s study cannot be directly mapped onto the close- and distant-PLD20 conditions in the present study. Recall that here we used word stimuli constrained to have *no* *N*-metric neighbors. This suggests that the PLD20 conditions in the present study are more appropriately evaluated against only the *sparse* neighborhood condition used previously by Goh and colleagues. As Fig. 2 shows, the frequency effect for sparse neighborhood words (54 ms) in the top panel resembles the average frequency effect for close- and distant-PLD20 words (54 ms) in the bottom panel. As such, the overadditive interaction between PLD20 and word frequency in the present study likely reflects processes that operate in the absence of competition from *N*-metric neighbors. Returning to the issue of the underadditive interaction between neighborhood density and word frequency observed by Goh et al. and by Luce and Pisoni, this is actually a surprising pattern, given that the effects of variables are almost always stronger for slower, more difficult stimuli. To our knowledge, neither Luce and Pisoni nor anyone else has highlighted or explained this intriguing pattern. Goh et al. provided some insights into this puzzle—specifically, by proposing that word frequency exerts both early (lexical access) and late (postlexical decision) effects on auditory lexical decision performance. Although it influences both early and late processes when words have sparse neighborhoods, frequency affects *only* the late process when words have dense neighborhoods, explaining the much larger frequency effect found previously for low-density words. Given that word frequency is theorized to influence both early and late processes when words have no neighbors (the present study), the underadditive pattern observed previously, when neighborhood density was manipulated, may not be relevant here.

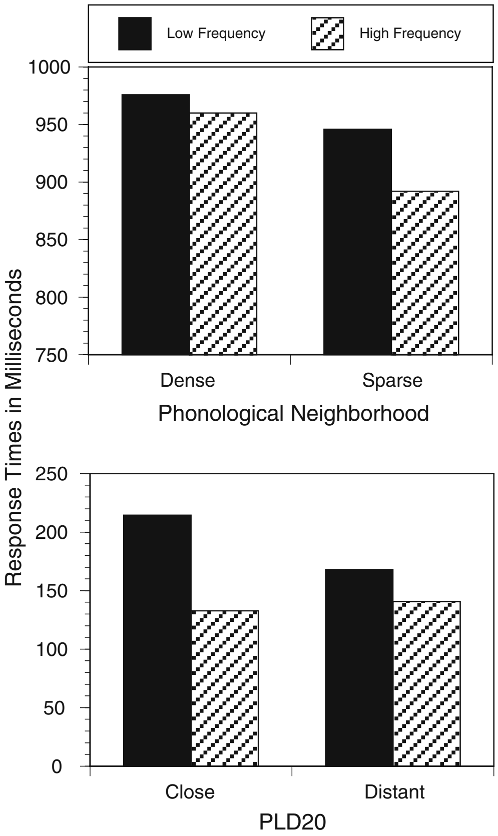

Fig. 2 Latencies in Goh et al. (2009; top panel) and the present study (bottom panel). Since the stimuli in the present study had no *N*-metric neighbors, frequency effects are more analogous to the sparse phonological neighborhood condition in Goh et al.'s study

The results from our study show that PLD20 predicts word recognition latencies in a way that is similar to traditional measures of phonological similarity. Indeed, Yarkoni et al. (2008) found that for short monosyllabic words, OLD20 and orthographic neighborhood size are very highly correlated and account for similar variances in visual word recognition. A similar observation can be made for PLD20 and phonological neighborhood density in spoken word recognition. The major contribution of PLD20, therefore, is that it has predictive power for longer multisyllabic words that have few or no traditional neighbors, in turn revealing potential statistical interactions between word-form similarity and other lexical variables in long words. Indeed, the present study makes it clear that competition from similar word forms can be robustly observed for lexical hermits, which possess no neighbors by all extant definitions of word-form similarity.

Author note This work was supported by Research Support Scheme C-581-000-222-091 to W.D.G. We thank Mike Vitevitch for the uniqueness point values, Tan Chin Hong for helping with data collection, and J. David Smith, Peter Richtsmeier, Steve Goldinger, and another reviewer for constructive comments on earlier drafts. L.S. is now at James Cook University.

Appendix

Words

Close PLD20, low frequency

aggravate
apparel
artery
aspirin
buffalo
canopy
condense
condiment
crucible
cuisine
desolate
domino
elevate
filament
hamster
harpoon
innovate
instigate
italic
koala
mammoth
marital
ominous
ornament
pageant
picnic
pinnacle
piranha
pretzel
renovate
scallop
sprinkle
squabble
stampede
stoic
sublime
tremble
twinkle

Close PLD20, high frequency

abandon
attract
banana
canteen
condition
consider
contact
cricket
fallible
fascinate
feasible
harness
opponent
parallel
particle
partner
sardine
segment
stereo
studio

carpet
cavity
clinic
clumsy
deliver
determine
dimension
discover
improve
intact
mineral
miracle
prosper
remember
respect
romantic
understand
violin

Distant PLD20, low frequency

amiable
apricot
bombard
calibrate
canary
carnation
carnival
charcoal
coleslaw
combust
concave
farewell
fiancee
fragile
gelatin
jasmine
karate
lacerate
maniac
membrane
merchandise
microbe
monsoon
onslaught
ostrich
pancreas
papaya
paragon
perfume
persevere
punctual
rendition
solicit
sucrose
suffocate
tranquil
turquoise
urchin

Distant PLD20, high frequency

abolish
brochure
challenge
chocolate
coalition
concrete
cutlery
decibel
develop
dissipate
electrode
example
grotesque
horizon
interview
jaguar
library
malleable
mischief
mosquito
museum
mutual
nephew
occupy
perhaps
platform
plausible
prestige
principle
rationale
reservoir
schedule
suggest
syndrome
technique
threshold
underneath
wardrobe

Nonwords

æbdɪkeɪl
ɛvəlɑsp
æmbjuskeɪl
æmpəsælt
ænsaɪ
əproʊpɪə
əreɪk
ɑtɪfænd
æsfɔst
ɒdu
bɔldədænd
bætəroʊ
bɪsip
buʒwi
brændɪp
brɪgəntɪʃ
braɪdəm
bʌmpoʊ
kæpʃəf
sɛntʃʊərɪəp
səribrəʃ
tʃeɪmbi
tʃænsəm
tʃɪmpænzu
klændɛstik
kɒmɪdɑ
kɒŋkleɪm
kɒnsɒmaɪ
kɒntjuʒəs
kɔdʒɪteɪn
kɔsɑθ
krəteɪʃəl
krɒkədaɪm
saɪklətrɒp
dændɪlaɪəs
dɪlɪʃəl
dɛnɪgreɪv
dɪtɜdʒəst
daɪətraɪl
dʌblus
daɪəsid
ɪklɛktɪm
ɪklɪptɪs
ɪvɛnk
ɪkspli
ɪgzɑlf
ətrɪʃəd
ɪskeɪn
ɛkskwɪzɪv
fɛldspu
flɑʊndju
fɔmæf
fræzəθ
frɪkəsu
fjusələð
gɑgæntʃuəf
dʒɛnjʊflɛʃ
dʒɛntid
glɒkənspik
gɒbləd
grɒtə
gʌmbɔ
hɑpsɪkɔk
haɪəsɪŋ
ɪmpɑm
ɪmpɪns
ɪmprɪsaɪp
ɪnfu
dʒʌmbət
kəbɑni
kəpril
kɛrəzip
kɪbʊtsɪk
kwɪnɪs
lɑdʒɛl
laɪtmoʊtik
liətɑf
lɛprɪkɔt
loʊkeɪm
lɒgdʒæz
meɪlstrəl
mælfizəst
mændɪbən
meɪnoʊ
mɛdsəp
mɜkəntaɪs
mɛθədʒ
mɛtrənoʊd
mɪldaɪ
mɪljɔ
mɒntɑθ
mɒrɪbʌnk
naɪik
nɛglɪʒɔ
nɪŋkəmpul
nɒnpərɛdʒ
nɒnplʌk
ɒnsɒmbəs
pælɪndroʊp
pænzaɪ
pəlɛmɪf
pəraɪoʊ
pəfjuʒəs
pəroʊkɪəm
pɜsənɛk
piæni
pɪstəθ
plɑstaɪ
poʊlkɔ
pɒntuz
prɪsɪʒəθ
prɛstaɪ
pʌŋktʃəs
pɜpəʃ
kwɒlɪfoʊ
ræmbʌŋkʃəd
ræproʊʃməst
rɪkʌvɪə
rɒndeɪvɪə
rɪpɑdʒ
rɛvərɔ
rɪgməroʊd
raɪvəlrɔ
rudɪmɛlt
sæŋgwɪnɪəθ
skɪtsɔɪn
skrʌmpʃəl
sɛntru
ʃændəlu
smɪðərist
smɔgəsbɔk
sɒlstɪk
spɑkət
skwɪgən
steɪməl
stɪgmɑtu
stɒlɪp
strɪknib
sʌksɪst
sʌltəf
sʌmptʃuəθ
swɒdən
tændʒəriz
tɛmpətʃ
θrɒtədʒ
tɔtijoʊ
træmpəlik
ʌmbrɪk
jʊrəθeɪtʃ
vɑsɪtɔ
vɛstɪbjun
vɪnjɛʃ

References

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review, 4*, 439–461.

Bailey, T. M., & Hahn, U. (2001). Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language, 44*, 568–591.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., et al. (2007). The English Lexicon Project. *Behavior Research Methods, 39*, 445–459.

Chan, K. Y., & Vitevitch, M. S. (2009). The influence of the phonological neighborhood clustering coefficient on spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 1934–1949.

Goh, W. D., Suárez, L., Yap, M. J., & Tan, S. H. (2009). Distributional analyses in auditory lexical decision: Neighborhood density and word frequency effects. *Psychonomic Bulletin & Review, 16*, 882–887.

Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language, 28*, 501–518.

Goldinger, S. D., Luce, P. A., Pisoni, D. B., & Marcario, J. K. (1992). Form-based priming in spoken word recognition: The roles of competition and bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 1211–1238.

Luce, P. A. (1986). A computational analysis of uniqueness points in auditory word recognition. *Perception & Psychophysics, 39*, 155–158.

Luce, P. A., Goldinger, S. D., Auer, E. T., & Vitevitch, M. S. (2000). Phonetic priming, neighborhood activation, and PARSYN. *Perception & Psychophysics, 62*, 615–625.

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

Magnuson, J. S., Dixon, J. A., Tanenhaus, M. K., & Aslin, R. N. (2007). The dynamics of lexical competition during spoken word recognition. *Cognitive Science, 31*, 133–156.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh: Psychology Software Tools, Inc.

Vitevitch, M. S. (2007). The spread of the phonological neighborhood influences spoken word recognition. *Memory & Cognition, 35*, 166–175.

Vitevitch, M. S. (2008). What can graph theory tell us about word learning and lexical retrieval? *Journal of Speech, Language, and Hearing Research, 51*, 408–422.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language, 40*, 374–408.

Vitevitch, M. S., Stamer, M. K., & Sereno, J. A. (2008). Word length and lexical competition: Longer is the same as shorter. *Language and Speech, 51*, 361–383.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature, 393*, 409–410.

Yap, M. J., & Balota, D. (2009). Visual word recognition in multisyllabic words. *Journal of Memory and Language, 60*, 502–529.

Yarkoni, T., Balota, D., & Yap, M. J. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review, 15*, 971–979.

Ziegler, J. C., Muneaux, M., & Grainger, J. (2003). Neighborhood effects in auditory word recognition: Phonological competition and orthographic facilitation. *Journal of Memory and Language, 48*, 779–793.