Word frequency effects in picture naming:

Which frequency measure to use for homophones?

Petroula Mousikou [1] Marc Brysbaert [2]

[1] Royal Holloway, University of London

[2] Ghent University

Keywords: word frequency, picture naming, homophones

Address: Petroula Mousikou

Department of Psychology

Royal Holloway, University of London

Wolfson building

Egham, Surrey, TW20 0EX, United Kingdom

Tel: +44 1784 414390

Email: Betty.Mousikou@rhul.ac.uk

Abstract

The present study investigated whether total word frequency or noun frequency better predicts production latencies in picture and word naming. We re-analysed two published databases, one in English and one in Dutch, and found that noun frequency is a better predictor of picture naming latencies, whereas total word frequency better predicts word reading times. Our results shed light on the theoretical debate about how homophones are represented in speech production, supporting the view that each word has a separate phonological representation (Caramazza, 1997). Critically, our findings are useful for researchers in the speech production domain as they indicate that the frequency measure is task dependent: when a picture (e.g., of a *saw*) is to be named, the frequency of the depicted entity must be used (i.e. the noun frequency of the word *saw*) and not the total word frequency (i.e. the frequency of the noun *saw* plus the frequency of the verb *saw*).

"Mine is a long and sad *tale*!" said the Mouse, turning to Alice, and sighing. "It is a long *tail*, certainly,"' said Alice, looking down with wonder at the Mouse's tail; "but why do you call it sad?" (Lewis Carroll, *Alice's adventures in Wonderland*). Homophones are words that are pronounced the same but differ in meaning. Homophones may be heterographic, such as *tale* and *tail*, which are spelled differently, or homographic, such as *saw* (tool) and *saw* (past tense of *see*). How homophones are represented and accessed in speech production is a longstanding debate in psycholinguistic research.

Two main hypotheses have been proposed. According to the shared representation (SR) hypothesis, homophones share a common phonological representation (lexeme), yet they have different semantic and often different grammatical representations (lemmas), e.g., *saw* can be a noun or a verb (Cutting & Ferreira, 1999; Dell, 1990; Jescheniak & Levelt, 1994; Levelt, Roelofs, & Meyer, 1999). According to the independent representation (IR) hypothesis, each word, homophonic or not, has a separate phonological representation (Caramazza, 1997; Harley, 1999). Both hypotheses are schematically represented in Figures 1A and 1B.

- -

Insert Figures 1A and 1B about here

- -

The vast majority of empirical studies in the literature examined the two hypotheses using word frequency as the independent variable. For example in an error elicitation task, Dell (1990) found that low-frequency (LF) words that were homophonous with high-frequency (HF) words (e.g., ***nun***-*none*) were no more prone to be pronounced incorrectly than their HF twins, suggesting that a LF homophone benefits from sharing its lexeme with a HF word. Using a different paradigm, Jescheniak and Levelt (1994) asked bilingual Dutch-English speakers to translate English words into their Dutch equivalent as fast as possible. Production latencies for Dutch LF homophonic translations with a HF twin were significantly faster than those for matched LF non-homophonic translations (see

Jescheniak, Meyer, & Levelt, 2003, for a similar result obtained with English-German bilinguals). Additionally, production latencies for LF homophonic translations were similar to those for non-homophonic HF translations matched on the cumulative frequency of the two homophonic forms. The results from both studies favored thus the SR hypothesis.

In contrast, in a picture-naming task carried out with both English and Chinese speakers, Caramazza, Costa, Miozzo, and Bi (2001) found that LF homophones (e.g., *nun*) were named as fast as word-specific LF controls (e.g., *owl*) and significantly slower than HF controls (e.g., *tooth*). In an attempt to replicate Jescheniak and Levelt's (1994) findings, a translation experiment using English-Spanish bilinguals and a larger set of stimuli was carried out in addition. The results indicated that word-specific and not homophone frequency determined production latencies, favoring the IR hypothesis.

Several other studies subsequently supported the finding that picture naming times depend on word-specific frequencies and not on combined homophonic frequencies. Bonin and Fayol (2002) compared French speakers' production latencies for the LF (e.g., *ver = worm*) and the HF (e.g., *verre = glass*) member of heterographic homophones (/ver/) in written and spoken naming. The HF member of the homophone pair (*verre*) was named much faster than the LF member (*ver*).

Miozzo and Caramazza (2005) investigated interference effects when homophone words were presented as distractors. A homophone's word-specific frequency proved to be a better predictor of its effect on picture naming than the combined phonological frequency. In Experiment 1 the stimuli consisted of LF homophones with a high cumulative frequency, such as *brake*, and two controls: a LF word matched to the presented word *brake* (e.g., *chord*) and a high frequency word matched to the cumulative frequency of *brake* and *break* (e.g., *offer*). Naming latencies to the pictures in the homophone contexts were similar to those of the LF controls. In a second experiment, the higher frequency counterpart of the homophones in Experiment 1 (e.g., *break*) was used, but the control conditions remained the same. Interference effects of the HF homophones were similar now to those observed with the HF controls.

More recently, Cuetos, Bonin, Alameda, and Caramazza (2010) asked Spanish and French participants to name pictures of HF and LF homophones and control pictures of non-homophone words matched in frequency with each of the two uses of the homophones. Naming latencies for LF homophones were longer than those for HF homophones and naming latencies for homophones were indistinguishable from those for non-homophone controls matched on word-specific frequency. Critically, in an object decision and a picture–word matching task using the same stimuli there were no reliable differences between HF and LF homophones, suggesting that the word-specific and not the cumulative homophone frequency determined lexical access in speech production.

Some neuropsychological studies have also examined the homophone frequency effect using aphasic speakers (e.g., Miozzo, Jacobs, & Singer, 2004). In this study, aphasics' naming performance on LF and HF homophones was compared to their naming performance on LF and HF controls. Response accuracy for LF homophones was similar to that for LF controls. Also, HF homophones were named as accurately as HF controls, indicating that a homophone behaves according to its word-specific frequency. However, a different set of neuropsychological studies (Biedermann, Blanken, & Nickels, 2002; Biedermann & Nickels, 1998a; 1998b) that specifically avoided to use word frequency as the independent variable in examining aphasics' picture naming performance, found different results. Namely, in these studies a member of a homophone pair was primed by providing phonological information when the picture could not be named (e.g., by giving the number of syllables or the initial phoneme). Despite the fact that aphasic participants were only exposed to pictures of one homophone meaning, at the end of the training they showed improved performance for both the trained and the untrained homophone pictures, supporting the idea that homophones share a common phonological representation.

The discrepancies between the findings of these studies are hard to reconcile using existing empirical and neuropsychological evidence. Yet, the SR and IR hypotheses make two straightforward predictions which we sought to test from a different perspective in order to provide fresh evidence to the theoretical debate on homophone representation in speech production. According to the SR hypothesis, naming latencies to pictures should be affected by the summed frequency of homophone words (i.e. frequency of *saw* as noun plus frequency of *saw* as verb). However, according to the IR hypothesis, noun frequency should be the best predictor of picture naming latencies. We tested these ideas using the new PoS-dependent frequencies in the SUBTLEX databases (Brysbaert, New, & Keuleers, 2012; Keuleers, Brysbaert, & New, 2010) which allowed us to compare noun-specific to total word frequencies as predictors of picture and word naming times. Word naming times served as a control given that there is no a priori reason why these should be influenced by the noun-specific frequency but not by the total frequency of the printed word forms.

To summarize, the primary aim of the present study is to determine whether noun-specific or total word frequency influences picture naming latencies. Our findings are critical for researchers of speech production and for advancing theory in this domain as the literature clearly indicates that inferences about how impaired and unimpaired speakers represent homophones have been mainly made on the basis of picture naming and other speech production studies that used word frequency as the independent variable.

Picture and word naming in English

The picture naming latencies from the International Picture Naming Project (IPNP) database (Szekely et al., 2003; 2004) were included in the analysis. This set contains pictures of 520 objects which can be downloaded from http://crl.ucsd.edu/experiments/ipnp/. Of these pictures, we excluded 38 because their names consisted of two-word names (washing machine, can opener, fire hydrant, etc.) or because they had no entry in the English Lexicon Project (ELP), a database with word naming latencies for over 40,000 English words (Balota et al., 2007; available at http://elexicon.wustl.edu/). Three frequency measures were extracted from SUBTLEX-US (Brysbaert et al., 2012; available at http://expsy.ugent.be/SUBTLEXus/): (1) The number of times the word is observed in the corpus (a measure called FREQcount), (2) the number of times the word is observed as a noun, and (3) the number of times the word is observed as any part-of-speech other than noun. All frequencies were log transformed after adding 1 to the frequency tally (Brysbaert & Diependaele, in press).

The correlations of the picture and word naming latencies with the various frequency measures are shown in Table 1. All correlations are highly significant except between picture naming and SUBTLEX-US_other. To determine whether one correlation is significantly higher than the other, we used the Hotelling-Williams test, because the variables are not independent (Steiger, 1980; see also http://crr.ugent.be/archives/546, retrieved on February 17, 2013). For this test, one must know the correlation between the frequency measures, which are also shown in Table 1. For picture naming times, the correlation with SUBTLEX-US_noun is higher than the correlation with SUBTLEX-US_all ($t(479) = 3.37$, $p < .001$). The difference between SUBTLEX-US_all and SUBTLEX_US_noun was not significant for word naming times ($t(479) = -.65$, n.s.).

- - - - - - - - - - - - - - - - - - - -

Insert Table 1 about here

- - - - - - - - - - - - - - - - - - - -

To further flesh out the results, we ran multiple non-linear regression analyses with restricted cubic splines (using the rcs function in R, with the minimum of three knots) as the frequency effect is known to be nonlinear (Baayen, Feldman, & Schreuder, 2006; Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Keuleers, Diependaele, & Brysbaert, 2010; Keuleers, Lacey, Rastle, & Brysbaert, 2012). These are the results (the R^2 values give the percentages of variance explained by the model):

Word naming time ~ rcs(SUBTLEX-US_noun,3)*** $R^2 = .309$

Word naming time ~ rcs(SUBTLEX-US_all,3)*** $R^2 = .336$

Word naming time ~ rcs(SUBTLEX-US_noun,3)*** + SUBTLEX-US_other*** $R^2 = .337$

Picture naming time ~ rcs(SUBTLEX-US_noun,3)*** R^2 = .206

Picture naming time ~ rcs(SUBTLEX-US_all,3)*** R^2 = .184

Picture naming time ~ rcs(SUBTLEX-US_noun,3)***+ SUBTLEX-US_other R^2 = .212

***p< .001

These analyses confirm the correlational analyses by showing that the frequencies from parts of speech other than nouns have a significant impact on word naming times (comparison of model with noun frequency only vs. noun and other frequency: $F(1,478) = 21.00$), but not on picture naming times (comparison: $F(1,478) = 3.45$). As a result, in the English language SUBTLEX-US_all is the predictor of choice for word naming times, whereas SUBTLEX-US_noun is the best predictor for picture naming times.

Picture naming in Dutch

To make sure that the findings with the IPNP dataset generalize to other datasets, we conducted the same analyses using picture naming data from a study conducted in Dutch (Severens, Van Lommel, Ratinckx, & Hartsuiker, 2005). In that study 590 pictures of objects were selected, yet one of them could not be used as it turned out to be unknown to the majority of the participants. Unfortunately, for the Dutch language there is no megastudy of word naming times. Therefore, our analysis is limited to picture naming times. The results from the comparisons of the SUBTLEX-NL frequencies (Keuleers et al., 2010) are shown in Table 2. As with the English data, the correlation between SUBTLEX-NL_noun and picture naming time was higher than the correlation between SUBTLEX-NL_all and picture naming time ($t(586) = 3.18$, $p < .002$; correlation between SUBTLEX-NL_all and SUBTLEX-NL_noun = .910). The superiority remained in nonlinear regressions with restricted cubic splines.

- - - - - - - - - - - - - - - - - - - -

Insert Table 2 about here

- - - - - - - - - - - - - - - - - - - -

Discussion

The vast majority of empirical and neuropsychological studies that examined how normal and aphasic speakers represent homophones in their mental lexicon used word frequency as the independent variable in speech production tasks (e.g., Caramazza et al., 2001; Jescheniak & Levelt, 1994; Miozzo et al., 2004). The results from these studies are inconsistent, with some of them providing evidence in favor of a common phonological representation for homophones (see Figure 1A) and the remaining supporting separate phonological representations (see Figure 1B). The present study sought to provide further evidence to this theoretical debate by investigating word frequency effects on picture and word naming latencies using data from large-scale databases.

Our results showed that noun-specific frequency is the best predictor of picture naming times, whereas total word frequency better predicts word reading times. This was confirmed in two languages, English and Dutch.[1] These findings are in line with the idea that each word has an independent phonological representation (Caramazza, 1997), contributing thus to the accumulating evidence for the IR hypothesis (Bonin & Fayol, 2002; Caramazza et al., 2001; Cuetos et al., 2010; Miozzo & Caramazza, 2005).

Critically, our findings have important methodological implications for researchers using the picture naming task, as they clearly indicate that when frequency is manipulated in a picture naming study or, even more importantly, when a set of picture stimuli need to be matched on frequency, noun

[1] There are two more picture naming databases in French (Bonin et al., 2003; Alario et al., 2004). However, we could not use them for further verification, as the picture stimuli in those databases were selected such that they barely included any homophone names (correlation between noun and total word frequencies in both databases was .99).

frequencies rather than (or in addition to) total word frequencies should be used. For example, even though the word “*saw*” is among the 300 most common words in English, the same is not true for the object “*saw*” and its name. Therefore, putting the picture of a “*saw*” in the high-frequency condition or matching it to another high-frequency picture name (such as “*sun*”) may yield misleading results.

The various SUBTLEX frequencies we used for IPNP and the Dutch picture dataset are included as supplementary materials to this article so that researchers using the picture naming task can use the noun-specific frequencies and check how to calculate the values for new materials. Another advantage of the SUBTLEX frequency measures is that they are considerably better than the Celex frequencies (Baayen, Piepenbrock, & Gulikers, 1995) currently provided in the IPNP database, which correlate only -.34 with the picture naming times.

References

Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U.,& Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, *36*, 140-155.

Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language*, *55*, 290–313.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (CD-ROM). Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*, 283-316.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., ...& Treiman, R. (2007). TheEnglish Lexicon Project. *Behavior Research Methods, 39*, 445–459.

Biedermann, B., Blanken, G., & Nickels, L. A. (2002). The representation of homophones: Evidence from remediation. *Aphasiology*, *16*, 935–980.

Biedermann, B., & Nickels, L. (2008a). The representation of homophones: More evidence from the remediation of anomia. *Cortex*, *44*, 276–293.

Biedermann, B., & Nickels, L. (2008b). Homographic and heterographic homophones in speech production: Does orthography matter? *Cortex*, *44*, 683–697.

Bonin, P., & Fayol, M. (2002). Frequency effects in the written and spoken production of homophonic picture names. *European Journal of Cognitive Psychology*, *14*, 289–313.

Bonin, P., Méot, A., Aubert, L., Malardier, N., Niedenthal, P., & Cappell-Toczek, M.-C. (2003). Normes de concrétude, de valeurd’imagerie, de fréquence subjective et de valence émotionnelle pour866 mots. *Année Psychologique*, *104*, 655-694.

Brysbaert, M., & Diependaele, K. (in press). Dealing with zero word frequencies: A review of the existing rules of thumb and a suggestion for an evidence-based choice. *Behavior Research Methods*.

Brysbaert, M., New, B., & Keuleers, E. (2012). Adding Part of Speech information to the SUBTLEX-US word frequencies. *Behavior Research Methods, 44*, 991-997.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive*

Neuropsychology, 14, 177-208.

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The specific-word frequency effect: Implications for the representation of homophones in speech production.*Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1430–1450.

Cuetos, F., Bonin, P., Alameda, J. R., & Caramazza, A. (2010). The specific-word frequency effect in speech production: Evidence from Spanish and French. *Quarterly Journal of Experimental Psychology, 63*, 750-771.

Cutting, J. C, & Ferreira, V. S. (1999). Semantic and phonological information in the production lexicon. *Journal of Experimental Psychology:Learning, Memory, and Cognition, 25,* 318-344.

Dell, G. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language & Cognitive Processes, 5,* 313–349.

Harley, T. A. (1999). Will one stage and no feedback suffice in lexicalization? *Behavioral and Brain Sciences, 22,* 45.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 824–843.

Jescheniak, J. D., Meyer, A. S., & Levelt, W. J. M. (2003). Specific-word frequency is not all that counts in speech production: Comments on Caramazza, Costa, et al. (2001) and new experimental data. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 432–438.

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new frequency measure for Dutch words based on film subtitles. *Behavior Research Methods*, *42*, 643-650.

Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice effects in large-scale visual word recognition studies: A lexical decision study on 14,000 Dutch mono- and disyllabic words and nonwords. *Frontiers in Psychology,* 1:174.

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods*, *44*, 287-304.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999).A theory of lexical access in speech production.

Brain and Behavioral Sciences, 22, 1–75.

Miozzo, M., & Caramazza, A. (2005). The representation of homophones: Evidence from the distractorfrequency effect. *Journal of Experimental Psychology: Language, Memory, and Cognition, 31,* 1360–1371.

Miozzo, M., Jacobs, M., & Singer, N. (2004). The representation of homophones: Evidence from anomia. *Cognitive Neuropsychology*, *21*, 840–866.

Severens, E., Van Lommel, S., Ratinckx, E., & Hartsuiker R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. Acta Psychologica, *119*, 159–187.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87*, 245-251.

Szekely, A.,D'Amico, S., Devescovi, A.,Federmeier, K., Herron, D., Iyer, G., Jacobsen, T., & Bates, E. (2003). Timed picture naming: Extended norms and validation against previous studies. *Behavior Research Methods*, *35*, 621-633.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D.,…& Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language*, *51*, 247-250.

Table 1.Correlations of word frequencies with word and picture naming times in the English language.

	SUBTLEX-US_noun	SUBTLEX-US_other	Picture naming	Word naming
SUBTLEX-US_all	.934**	.521**	-.394**	-.541**
SUBTLEX-US_noun		.318**	-.444**	-.532**
SUBTLEX-US_other			-.072	-.335**
Picture naming				.297**

N = 482, ** p < .001

Table 2.Correlations of word frequencies with picture naming times in the Dutch language.

	SUBTLEX-NL_all	SUBTLEX-NL_noun	SUBTLEX-NL_other
Picture naming	-.305**	-.357**	-.097*

N = 589, * p < .02, ** p < .001

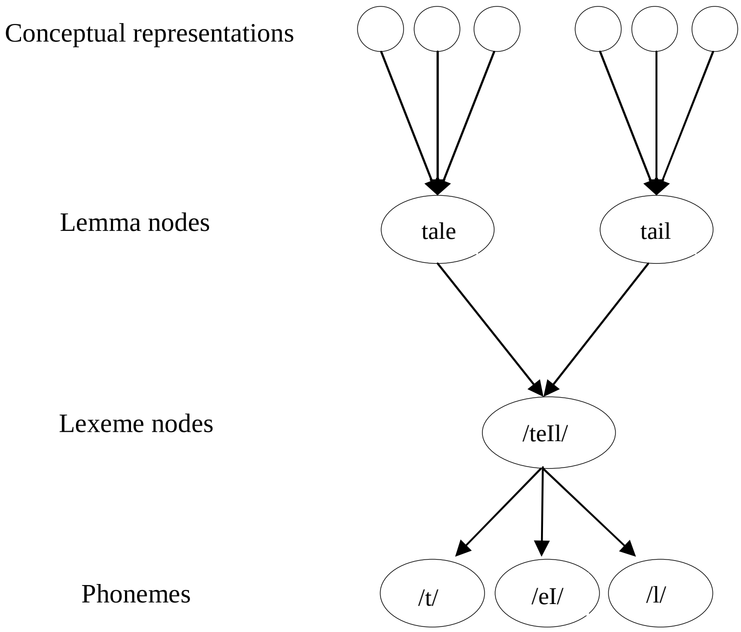

Figure 1A. Shared representation hypothesis

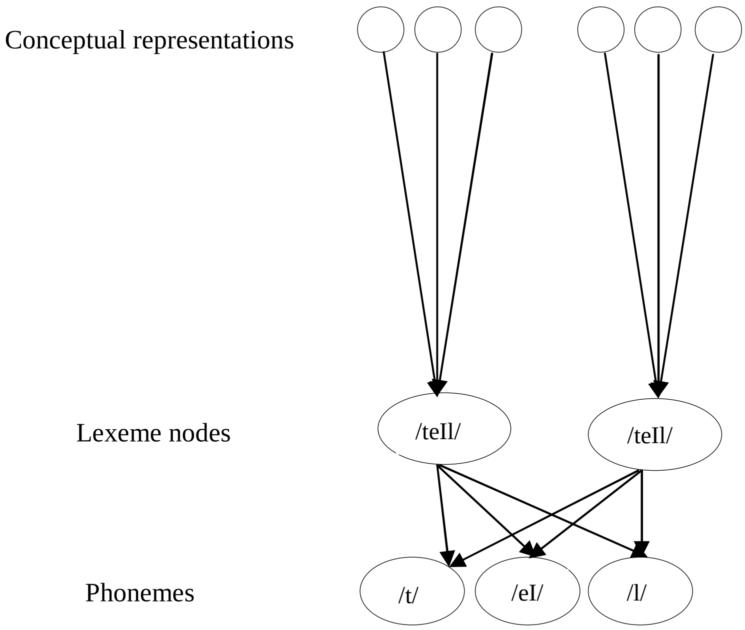

Figure 1B. Independent representation hypothesis