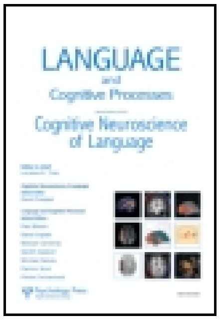

Language and Cognitive Processes
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/plcp20

The ageing neighbourhood: phonological density in naming

Jean K. Gordon[abc] & Jake C. Kurczek[b]

[a] Communication Sciences & Disorders, University of Iowa, 305 WJSHC, Iowa City, IA 52242-1012, USA

[b] Neuroscience Graduate Program, University of Iowa, 2100 RCP, Iowa City, IA 52242-1101, USA

[c] DeLTA Center, University of Iowa, Iowa City, IA 52242, USA

Published online: 20 Sep 2013.

Language, Cognition and Neuroscience, 2013
http://dx.doi.org/10.1080/01690965.2013.837495

The ageing neighbourhood: phonological density in naming

Jean K. Gordon[a,b,c]* and Jake C. Kurczek[b]

[a]*Communication Sciences & Disorders, University of Iowa, 305 WJSHC, Iowa City, IA 52242-1012, USA;* [b]*Neuroscience Graduate Program, University of Iowa, 2100 RCP, Iowa City, IA 52242-1101, USA;* [c]*DeLTA Center, University of Iowa, Iowa City, IA 52242, USA*

(*Received 7 February 2012; accepted 14 August 2013*)

Ageing affects the ability to retrieve words for production, despite maintenance of lexical knowledge. In this study, we investigate the influence of lexical variables on picture naming accuracy and latency in adults ranging in age from 22 to 86 years. In particular, we explored the influence of phonological neighbourhood density, which has been shown to exert competitive effects on word recognition, but to facilitate word production, a finding with implications for models of the lexicon. Naming responses were slower and less accurate for older participants, as expected. Target frequency also played a strong role, with facilitative frequency effects becoming stronger with age. Neighbourhood density interacted with age, such that naming was slower for high-density than low-density items, but only for older subjects. Explaining this finding within an interactive activation model suggests that, as we age, the ability of activated neighbours to facilitate target production diminishes, while their activation puts them in competition with the target.

Keywords: word retrieval; phonological neighbourhood density; ageing

Background

One of the more frustrating aspects of normal cognitive ageing is the inability to retrieve a known word accurately and efficiently. As we age, we encounter with increasing frequency the tip-of-the-tongue (ToT) feeling that James referred to as an "intensively active" gap (James, 1890), that is, the awareness of a word's existence, along with hints as to its identity, without the ability to grasp its form. Most often, the feeling is fleeting and the word form is retrieved, but the experience of even a brief delay in delivering the message can be unexpected and disconcerting.

Whether or not a word is successfully retrieved depends on many factors, including the integrity of the lexical access mechanisms and the general state of the cognitive system, which are almost inevitably affected by age. Retrieval success also depends on certain characteristics of the intended words relative to those with which they are competing for retrieval, and characteristics of the context in which they are being retrieved. Of particular interest is the influence on retrieval of a word's neighbourhood density (ND), or the number of words in the lexicon which are phonologically similar to the word being retrieved (Landauer & Streeter, 1973), because this variable has been shown to have apparently contradictory influences in different modalities of lexical processing, a finding with implications for models of the lexicon. In word recognition, the acoustic stimulus activates a number of phonologically similar potential word candidates. The larger this "neighbourhood" is, the more difficult it is to recognise a word (Luce & Pisoni, 1998). In word production, paradoxically, having more neighbours may make a word *easier* to produce (Gordon, 2002; Vitevitch, 2002). A number of hypotheses have been proposed to account for these apparently contradictory findings, but there has been little effort to reconcile them within the same explanatory framework (but see Chen & Mirman, 2012; Dell & Gordon, 2003).

In the current study, we analyse the phonological characteristics of words in the lexicon, how they influence lexical processing during production of words, and how these effects change with age. Our aims are to clarify the mechanisms by which phonological neighbours exert their influence on naming performance, and to examine how ageing affects these mechanisms.

Word retrieval in ageing

In experimental studies, typically using picture-naming tasks, older subjects – especially those over 70 years of age – are able to name pictures less quickly and/or accurately than younger subjects (e.g. Au et al., 1995; Connor, Spiro, Obler, & Albert, 2004; Feyereisen, 1997; Mortensen, Meyer, & Humphrey, 2006). Since older adults typically demonstrate more word knowledge on vocabulary tests than do younger adults (Brown, 1991; Dahlgren, 1998; James & Burke, 2000; Verhaeghen, 2003), this decline is accounted for, not by a shrinking

*Corresponding author. Email: jean-k-gordon@uiowa.edu

lexicon, but by a breakdown in the process of translating a known lexical concept into its word form. This breakdown often manifests itself as a ToT experience, which is characterised by a feeling of confidence in the word's existence without access to the word's full phonological form. In fact, older subjects have frequently been shown to experience more ToT states than younger subjects (Burke, Locantore, Austin, & Chae, 2004; Burke, MacKay, Worthey, & Wade, 1991; Evrard, 2002; Farrell & Abrams, 2011; Vitevitch & Sommers, 2003).

Burke and colleagues attribute this age-related decline to a "transmission deficit", in which retrieval is affected by a weakening of the connections between lexical and phonological nodes. To support this idea, they cite findings that phonological priming improves access to the targeted ToT word (Burke et al., 1991, 2004; James & Burke, 2000), and that older subjects often remember less information about the target word during ToT episodes than do younger subjects (Brown, 1991; Dahlgren, 1998). While priming entails an additional input to the lexicon in the form of the prime, words that are phonologically related to a target are also activated internally, through automatic spreading activation within the lexicon (Dell, 1986). The influence of these phonologically related words has been explored further in studies of phonological ND.

Phonological neighbourhood effects

The Neighbourhood Activation Model (Luce & Pisoni, 1998) was developed to account for findings by Luce and colleagues that recognition is slowed for words which must be distinguished from a larger number of phonologically similar competitors (Luce & Large, 2001; Luce & Pisoni, 1998; Vitevitch & Luce, 1998, 1999; Vitevitch, Luce, Pisoni, & Auer, 1999). In this model and since, ND has usually been operationally defined as the number of words which differ from the target by one phoneme, either added, substituted or deleted.

ND has also been shown to have an effect on speech production, but the effect in this modality appears to be facilitative, in both young adults (Harley & Bown, 1998; Vitevitch, 1997), and in adults with aphasia (Best, 1995; Goldrick, Folk, & Rapp, 2010; Gordon, 2002; Kittredge, Dell, Verkuilen, & Schwartz, 2008; Middleton & Schwartz, 2010). Although some of these studies, particularly earlier ones, did not fully compensate for the presence of confounded variables (e.g. Best, 1995; Goldrick et al., 2010; Gordon, 2002; Harley & Bown, 1998; Vitevitch, 1997), later studies attempted to address this problem either statistically (e.g. Middleton & Schwartz, 2010; Perez, 2007), or by analysing small sets of tightly controlled stimuli. Vitevitch (2002) demonstrated facilitative effects of density in picture naming by younger adults, even when other lexical variables (frequency, neighbourhood frequency, familiarity and phonotactic probability) were held relatively constant. Perez (2007) replicated this facilitative effect, this time in Spanish-speaking young adults. In a multiple regression analysis, phonological ND contributed significantly after measures of target frequency, age-of-acquisition and characteristics of the image. In a ToT elicitation study, Vitevitch and Sommers (2003) found that young adults produced more correct responses and fewer ToTs for words from dense than sparse neighbourhoods, although phonotactic probability was not controlled in this study.

However, density effects in word production are far from consistent. One of the issues complicating the study of neighbourhood effects is the fact that it is potentially confounded with so many other lexical variables, such as frequency of occurrence, length, and phonotactic probability. Attempting to tease apart the effects of ND and phonotactic probability, Vitevitch and colleagues (Vitevitch, Armbruster, & Chu, 2004) demonstrated that, when the two variables were orthogonally manipulated, only phonotactic probability showed significant effects on word production. They proposed that, according to the dynamics of speech production, "phonological segments should, in general, dominate processing and determine how quickly and accurately a word form would be produced" (p. 521).

In some contexts, competitive effects of ND on word production have been shown. In an effort to minimise the almost inevitable influence of confounded variables in real word stimuli, Frank and colleagues (Frank, Tanenhaus, Aslin, & Salverda, 2007) had participants learn novel names for shapes. Following training, they found that naming latencies were affected by both frequency and density, with high-frequency words named faster than low-frequency words, and low-density words faster than high-density words. Newman and German (2005) found an inhibitory density effect on naming accuracy. Adolescents, young and older adults all showed better naming for sparse than dense words. Task differences may have contributed to these opposing effects, since responses in this study were elicited through a variety of semantic inputs, including pictures, definitions and open-ended sentences, and measured accuracy rather than latency.

Vitevitch and Stamer (2006) found faster naming for Spanish words from sparse than dense neighbourhoods. Baus and colleagues (Baus, Costa, & Carreiras, 2008) provided evidence to suggest that this was an artefact of the picture stimuli used (a claim later disputed, Vitevitch & Stamer, 2009). A recent large-scale regression analysis of picture naming in Spanish (Sadat, Martin, Costa, & Alario, submitted) showed that young adults were slower to name items with more neighbours. Furthermore, the authors re-analysed several previous data sets, and demonstrated that previously reported facilitative effects (e.g. Baus et al., 2008) disappeared when subjected to regression techniques in order to control confounding variables. These studies leave open questions about the conditions under which density facilitation is exerted, and the role of related lexical characteristics.

Neighbourhood effects in ageing

Exploring the role that different lexical variables, particularly phonological ND, might play in moderating age-related word retrieval declines can inform models of lexical access. ND is of particular interest to study age-related decline in word retrieval, because it provides an opportunity to assess the influence of spreading activation within the lexicon, a process which has been implicated in ageing, as described above. Whereas variables such as length, lexical frequency and phonotactic probability are properties inherent to a given target word, ND is a characteristic of the target's lexical neighbourhood. Thus, examining ND allows us to examine the influence of co-activated items relative to the target item, given the properties (e.g. length, frequency) of the target item.

A number of studies have directly compared the effects of lexical variables on word retrieval by younger and older adults. Taler and colleagues (Taler, Aaron, Steinmetz, & Pisoni, 2010) found density facilitation in a repetition task, for both younger and older subjects, under the hypothesis that the duration of the response reflected speech-planning processes. Spieler and Balota (2000) also showed a facilitation effect of density, as measured by Coltheart's N, on "word naming" (i.e. oral reading) by younger and older subjects. However, they also noted a difference between the subject groups in the amount of naming latency variance accounted for by different variables, and proposed that, with age and its attendant experience with words, lexical variables like target frequency come to influence responses more than phonological variables like length and density, because words "gradually become compiled into more unitary representations rather than as assemblages of sublexical parts" (Spieler & Balota, 2000, p. 226). However, neither study controlled for phonotactic probability, which somewhat reduces our confidence in the attribution of the results to ND. Furthermore, the relevance of results from tasks like repetition and oral reading – which provide clues to phonology in the input – to word retrieval in functional speech production – in which the input is semantic – is unclear. Nevertheless, it is reasonable to hypothesise that Spieler and Balota's account of their findings that lexical representations become increasing "unitised" with age and language experience might also apply to picture naming.

Vitevitch and Sommers (2003) investigated the role of ND in production in younger and older speakers using a ToT paradigm. Although no overall age effect was demonstrated, an interaction was found between age, neighbourhood frequency and ND, such that younger subjects experienced more ToT states for words from sparse than dense neighbourhoods, regardless of the frequency of the words in the neighbourhood, but older subjects showed this density effect only for words from low-frequency neighbourhoods. Taken together, these findings suggest a facilitative (but somewhat volatile) effect of ND which appears to be reduced in older subjects, consistent with the account of Spieler and Balota (2000).

Modelling phonological neighbourhood effects

It makes sense that, in recognition, the denser a word's neighbourhood, the longer it takes to distinguish it from its many competitors. This is because word recognition is driven by the incoming acoustic-phonological information. Production, on the other hand, is primarily a semantically driven task. It is, therefore, less apparent why the phonological neighbourhood in which the target resides should influence lexical retrieval in speech production.

In interactive spreading activation (ISA) models of lexical access such as the one proposed by Dell and his colleagues (Dell, 1986; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997), phonological-to-lexical feedback accounts for the facilitative effect of ND in speech production. Activation from shared phonemes back to related lexical units sets up a reverberating pattern of activation that converges, under normal circumstances, on the target and reinforces it. Occasionally, the presence of noise or damage to the system can result in the accidental production of these simultaneously active words. For the most part, however, the presence of phonological neighbours facilitates correct production of the intended word. According to this account, the more words that are phonologically related to the target, the more activation feeds back from these items to the target, through their shared phonemes.

This principle was demonstrated in a simple computational model simulating lexical retrieval similar to the one illustrated in Figure 1 (Dell & Gordon, 2003; Gordon & Dell, 2001). Three neighbourhoods were constructed, in which the target word had either two phonological neighbours ("dense"), one neighbour ("sparse"), or no neighbours ("empty"). To simulate word retrieval in production, the set of semantic features corresponding to a target word was given a "jolt" of activation, which then spread throughout the network in both feed-forward and feedback directions for a specified number of time steps. At this point, the lexical node with the highest level of activation was determined to be the one retrieved. Although performing at accuracy levels close to ceiling, the model showed the predicted facilitative effects of ND in production: the dense network was most accurate, followed by the sparse then empty networks (see Table 1).

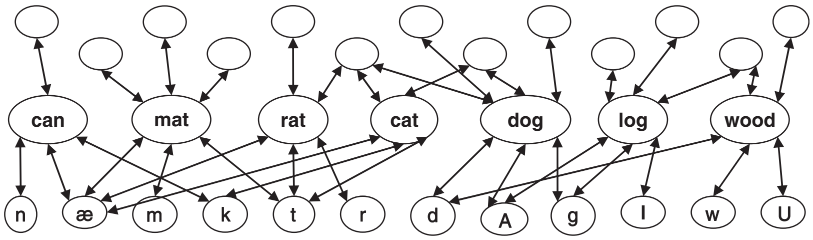

Figure 1. Example of a two-step IA model of lexical access.

When run in reverse (i.e. giving a jolt of activation to the phoneme nodes), the model demonstrated the predicted competitive effect: targets from the dense neighbourhood were recognised least accurately, followed by the sparse then empty neighbourhoods. Although highly simplified, the model demonstrates an important principle: lexical neighbours can exert quite different effects on word retrieval depending on the task, even when both tasks are simulated using the same stimuli, the same lexical architecture and the same processing mechanisms.

Modelling phonological neighbourhood effects in ageing

How might ageing affect this process? According to the Transmission Deficit Hypothesis, the mechanism by which neighbours are activated – activation spreading along connections between lexical representations and their component phonemes – weakens with age (Burke et al., 1991, 2004; Burke & Shafto, 2004). As noted above, this hypothesis has received considerable empirical support. To assess the impact of such weakened connections on neighbourhood effects, we appeal to an existing model of word production originally developed to account for lexical retrieval deficits in aphasia.[1] Dell and colleagues (Dell et al., 1997) successfully simulated different patterns of errors in individuals with aphasia by either increasing the rate at which activated representations decay or decreasing the connection weights among representations in the lexicon. In later simulations, Dell and Gordon (Dell & Gordon, 2003; Gordon & Dell, 2001) examined the influence of phonological neighbours in production by each of these damaged networks.

Results showed that different types of damage resulted in different effects of ND, as shown in Table 1. In the decay-lesioned model, facilitative ND effects on accuracy were preserved, although accuracy levels overall were reduced. Of greater relevance to the current study is the weight-lesioned model, as this provides a direct simulation of Burke's TDH. When connections were weakened, neighbourhood facilitation disappeared. In fact, *more* errors occurred in the dense neighbourhood than in the sparse or empty neighbourhoods, because the neighbours increased the opportunities for phonological neighbours to be produced. This leads to the prediction that, with age, phonological neighbourhood facilitation diminishes, contributing to lexical retrieval difficulties, and that this arises as a direct result of global weakening of lexical connections. In the current study, we test this prediction by examining the influence of phonological ND on naming in younger and older adults.

The current study

Although the effects of phonological ND in word production are widely cited to be facilitative, there have only been only a few studies of non-brain-damaged adults which show robust facilitation effects (most notably Perez, 2007; and Vitevitch, 2002, Experiment 4), and several which call this result into question (e.g. Frank et al., 2007; Newman & German, 2005; Vitevitch et al., 2004; Sadat et al., submitted). It is clear that one cause of these discrepancies is the confounding of ND with other lexical variables. However, addressing this issue in the typical way, by using small sets of tightly controlled stimuli (e.g. Garlock, Walley, & Metsala, 2001; Mirman, Kittredge, & Dell, 2010; Vitevitch, 2002, 2003; Vitevitch et al., 2004; Vitevitch & Rodriguez, 2005; Vitevitch & Stamer, 2006; Vitevitch, Stamer, & Sereno, 2008), may carry its own problems, as this accentuates the impact of item-specific

Table 1. Simulating neighbourhood effects in damaged lexical networks.

	Dense	Sparse	Empty
Undamaged network			
Overall accuracy (%)	97.8	97.6	96.9
Lemma accuracy (%)	97.9	97.8	97.6
Phonological accuracy (%)	99.9	99.8	99.3
Decay-lesioned network			
Overall accuracy (%)	47.8	40.2	25.4
Lemma accuracy (%)	66.6	66.0	65.0
Phonological accuracy (%)	71.7	60.9	39.1
Weight-lesioned network			
Overall accuracy (%)	33.7	35.1	36.4
Lemma accuracy (%)	62.5	65.6	68.7
Phonological accuracy (%)	54.0	53.4	52.9

Note: The undamaged network shows the facilitative effect of ND. Increasing decay rate reduced accuracy, but maintained facilitative effects of ND. Weakening connections reduced accuracy and reversed the ND effect. (Adapted from Dell & Gordon, 2003.)

stimulus characteristics, and may reduce the replicability of the findings (cf. Baus et al., 2008; Spieler & Balota, 2000; Vitevitch & Stamer, 2006).

An additional variable which may influence the effects of ND is the age of the subjects. The vast majority of studies to date have been conducted with young adults. In the few studies involving older subjects, there is evidence that the influence of phonological ND becomes attenuated with age (Spieler & Balota, 2000; Vitevitch & Sommers, 2003), and this is supported by modelling work (Dell & Gordon, 2003).

Neighbourhood effects in production, when found, are clearly susceptible to influences from other stimulus and task factors, and to the age of the participants. In the current study, we provide a well-controlled and ecologically valid assessment of the influence of phonological ND throughout adulthood by (a) using a large set of stimuli allowing the simultaneous analysis of several contributing variables (described below); (b) analysing lexical characteristics as continuous variables rather than dichotomising them, which results in a loss of power; (c) testing a group of adult participants ranging in age beyond the typically tested age range; (d) employing a statistical technique that takes into consideration all the data for individual participants and individual items, again to maximise power; and (e) analysing both accuracy and latency of word production. Our goals were to provide a methodologically rigorous test of density effects observed in younger adults; to examine how these effects are influenced by ageing; and to discuss the implications for models of lexical retrieval and ageing. We expected facilitative effects of ND for younger subjects, as predicted by interactive models of lexical production (Dell & Gordon, 2003; Gordon & Dell, 2001). In our older participants, we hypothesised that effects of ND would be reduced, as predicted by a weakening of lexical connections (Dell & Gordon, 2003; Gordon & Dell, 2001).

Methods

Study recruitment and procedures were carried out with the approval of the Institutional Review Board at the University of Iowa.

Participants

Non-brain-damaged participants between the ages of 21 and 90 years were recruited from the community in order to capture performance across a wide range of adulthood. All were native English speakers with no history by self-report, of language or learning disability, neurological or psychiatric problems. The resulting participant pool consisted of 73 individuals (45 women), ranging in age from 22 years to 86 years with a mean age of 51.8 (see Figure 2). To assess the distribution of demographic variables across our participant pool, participants were broken down into 3 age groups: 31 “Young-Middle” adults (YM, defined as those under 50 years of age); 22 “Middle-Old” adults (MO, those between 50 and 70 years of age) and 20 “Older-Old” adults (OO, those 70 or older). A one-way analysis of variance (ANOVA) demonstrated that these age groups did not differ significantly from each other in years of education ($F_{2,70}$, p = 0.388). Chi-square tests showed that, although the gender breakdown was unequal in all groups, none of the sub-groups differed from the larger group (62% women) in the relative proportions of men and women (YM: p = 0.951; MO: p = 0.607; OO: p = 0.485). Note that it is representative of the ageing population to include a larger proportion of women than men, as was the case in our sample. Demographic information about the recruited participants in each group is provided in Table 2.

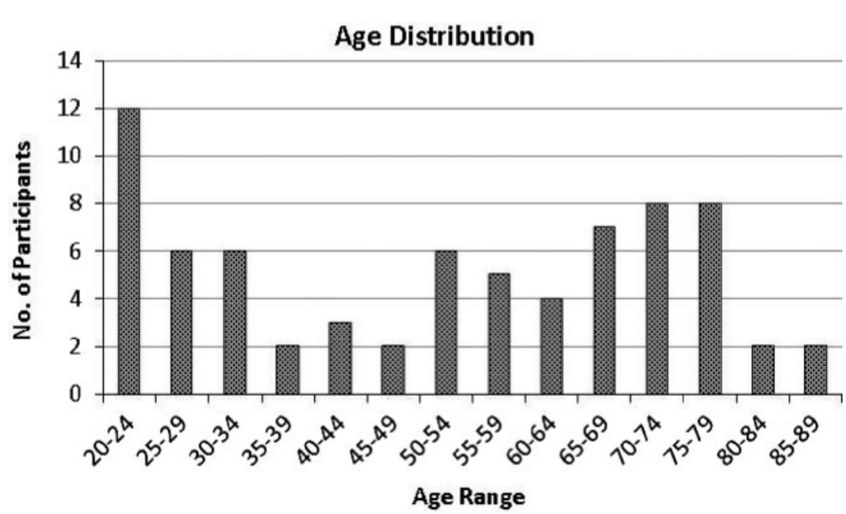

Figure 2. Distribution of participant ages.

Table 2. Participant demographics.

Participant group	N	Gender (% F)	Mean age (years)	Standard deviation (years)	Mean education (years)	Naming accuracy	Naming reaction time
Young-middle	31	0.613	30.2	7.7	18.0	0.975	903.5
Middle-old	22	0.591	59.7	6.2	18.9	0.962	999.5
Older-old	20	0.650	76.4	4.9	17.4	0.948	1031.5

Note: Groups were not significantly different in mean education level or gender breakdown. Naming accuracy was lower, and naming response time longer, for older participants.

Procedure

Participants named 400 black-and-white line drawings gathered from several sources: the *Boston Naming Test* (Kaplan, Goodglass, & Weintraub, 1978); the *Object & Action Naming Battery* (Druks & Masterson, 2000); the *Philadelphia Naming Test* (Roach, Schwartz, Martin, Grewal, & Brecher, 1996); and the Snodgrass-and-Vanderwart-like pictures generated by the TarrLab (Rossion & Pourtois, 2004). Target labels varied in length from one to three syllables; however, the analyses for this study include only the 200 single-syllable items from the test (see Appendix 1), in order to minimise the confounding effects of length on ND (Gordon, 2002). ND and neighbourhood frequency data were collected from Mitchell Sommers' online database of the Hoosier Mental Lexicon (http://128.252.27.56/neighborhood/Home.asp). Phonological probability statistics were collected from Michael Vitevitch's online probability calculator (http://www.people.ku.edu/~mvitevit/PhonoProbHome.html; Vitevitch & Luce, 2004). Frequency data were derived from the database of Francis and Kuçera (Francis & Kuçera, 1982). We also counted the length, in number of phonemes, for each item. The line drawings were digitised for computer presentation.

To familiarise the participants with the pictures and their intended labels, the pictures were presented to each participant in random order along with their written labels. During the subsequent naming task, the pictures were again presented, in a different random order, on a computer using E-Prime software (Schneider, Eschman, & Zuccolotto, 2002a, 2002b). Participants were asked to name the objects in one word as quickly as possible. Oral responses tripped a voice response box which allowed the software to log response latencies from the onset of the picture presentation to the onset of the response. A practice session was conducted first, both to familiarise the participant to the task, and to adjust the sensitivity of the voice response box tripping.

Responses were scored for accuracy online, and also digitally audio-recorded using a Shure SM10A headset microphone and a Marantz PMD 680 digital recorder. A different coder listened to the audio-recordings to double-check the online accuracy scores. Responses were considered accurate only if the intended target was produced as the first complete response.

Analyses

Variable selection

Because phonological ND was of primary interest in this study, we included this variable (defined as the number of words different from the target by one phoneme, either substituted, added or omitted) and the mean frequency of a target's neighbours (Sommers, http://128.252.27.56/neighborhood/Home.asp). In addition, we began with seven variables that are confounded with density measures, including one measure of length, three measures of phonotactic probability (Vitevitch & Luce, 2004) and three measures of lexical frequency (Francis & Kuçera, 1982): (1) number of phonemes (recall that all of the analysed stimuli were single-syllable); (2) initial phonotactic probability, the probability of the initial phoneme occurring in that position across words; (3) mean phonotactic probability, the average probability of occurrence of all the word's component phonemes; (4) mean biphone probability, the probability of occurrence of the word's biphones[2]; (5) frequency of the noun lemma (i.e. all inflectional forms, singular and plural, of the word used as a noun); (6) frequency of the noun lexeme (i.e. only the phonological form applying to the picture, usually singular, of the word used as a noun) and (7) total frequency of the lexeme (i.e. the phonological form of the word in all uses). All frequency variables were log-transformed to correct for positive skewness.

This list of variables was narrowed down using theoretical criteria (as follows), and by examining their intercorrelations and correlations with each of the outcome measures (see Appendix 2). On the basis of these explorations, we selected density and four other variables to include in our regression analyses, in order to minimise multicollinearity which can weaken the power of a regression analysis. Of the three phonotactic probability measures, initial phoneme probability (Initial.PP) was included because of the potential influence of the initial phoneme on the voice-activated measure of naming latency. Biphone probability (Mean BP) was selected to represent overall phonotactic frequency, because it was less closely related to density and initial phoneme probability than the mean phoneme probability. Length in phonemes (Phon) was retained in order to be able to directly examine the extent to which both density and length influence word retrieval. To deal with the virtually unavoidable covariance between length and density, we regressed density on length, and substituted the residuals for the density measure (see Kittredge et al., 2008; Storkel, 2004, for similar approaches). These residuals (Dens Resid) reflect the effect of ND beyond what would be predicted by length, and were strongly correlated with the original density measure (r = 0.767). Noun lemma frequency, log transformed (Log Lemma), was selected from among the lexical frequency measures, because it was considered to be the most relevant measure for a task involving naming of objects. It showed a strong correlation with the outcome variables, but a weaker relationship to density. Histograms showing the distributions of these lexical characteristics are shown in Appendix 3.

Data clean-up

Reaction times were analysed only for accurate (target) responses, resulting in the exclusion of 3.5% of items across participants. One item (rat) was excluded because it was named with an accuracy rate below the criterion of 50% (41% accurate) across participants, as this was an indication that the picture was a poor representation of the intended target (most subjects responded "mouse"). RTs were manually measured for trials in which the voice response box was not tripped or mis-tripped by extraneous noise (e.g. a tongue-click); however, 6% of mistripped responses were unrecoverable due to a problem with the audio recording. Reaction time data from two young participants (S1069 and S1070) were excluded altogether for exceeding the criterion of 30% mis-tripped items. Finally, a few responses (0.08%) were excluded as likely mis-trips, because they had RTs shorter than the minimum criterion of 400 msec. Overall, 4.3% of items (606/14,129) were excluded, leaving 90.6% automatically recorded response times, and an additional 5.1% manually recorded responses.[3]

Statistical analyses

Correlational analyses were conducted first to explore the structure of the data, and to help us narrow down our set of lexical variables to a manageable and meaningful set. Next, we conducted mixed regression analyses using R version 2.14.1 (R Development Core Team, 2011), and treating both participants and items treated as random factors (Baayen, Davidson, & Bates, 2008; Clark, 1973; Jaeger, 2008). These models have the advantage of allowing us to take advantage of the full range of data across individuals and across items, and to consider how subject and item variables interact. Fixed variables included the stimulus variables discussed below, as well as participant age and education level, both expressed in number of years. To determine whether the influence of stimulus characteristics on word retrieval changes with age, interactions of each lexical variable with age were also included in the models.

Results

Correlational analysis

Our initial correlational analyses confirmed our prediction that age was significantly correlated with mean outcome measures across items (Accuracy: $r = -0.436$; RT: $r = 0.347$, $p < 0.01$), although education was not (Accuracy: $r = -0.084$; RT: $r = -0.026$). The education variable was, therefore, excluded from further analyses. Of the lexical variables, only the frequency measures were significantly correlated with accuracy and RT, both across and within age groups. However, assessing the independent influences of each variable requires a regression analysis.

Regression analyses

Mean accuracy and reaction time measures for each of the age groups are shown in Table 2. It is evident that, overall, ageing results in a slight decrease in accuracy and a slowing of response latency. However, in order to capture the full range of variability, both in the participants and in the items, we conducted regression analyses treating participants and target items as random variables. For each outcome measure (accuracy and reaction time), we conducted a forward stepwise series of regression models, manually adding in fixed variables or groups of fixed variables at each step. For the accuracy analysis, we first ran a "Random" model, which included only the random effects of Subject and Item. Next, we ran an "Age" model, which included the fixed effect of Age added to the random variables. The "Main Effects" model was the Age model with the five retained lexical variables added (Length, Initial Phoneme Probability, Mean Biphone Probability, Density Residuals Regressed on Number of Phonemes, and Log-transformed Noun Lemma Frequency). Finally, the "Full" model included all of the above variables, as well as interaction effects of each lexical variable with age. For the reaction time analysis, RTs were transformed into z-scores in order to factor out individual differences in response speed. Because this manipulation effectively removed the main effect of age-related slowing, we skipped the comparison of Random and Age regression models in the successive model comparisons of reaction time. ANOVAs were run to determine whether each successive model fit the data significantly better than the preceding model, as indicated by a chi-squared test.

Accuracy of naming

Mixed logistic regression analyses were conducted on the accuracy data. Model comparisons showed that the Age

Table 3. Model estimates for naming accuracy.

Naming accuracy (Main Effects Model)					
Random effects	Variance	Std dev.			
Target	2.2030	1.4842			
Subject	0.3953	0.6287			
Fixed effects	Coefficient	Std error	*z*-Value	*p*-Value	Odds
Intercept	5.1910	0.7969	6.51	<0.001	179.648
Age	−0.0235	0.0044	−5.40	<0.001	0.977
Length (Phon)	−0.2727	0.2063	−1.32	0.186	*NS*
Initial.PP	5.5183	4.8220	1.14	0.252	*NS*
Mean.BP	21.0545	45.7800	0.46	0.646	*NS*
Dens.Resid	−0.0282	0.0183	−1.54	0.123	*NS*
Frequency (Log.Lemma)	0.2795	0.0920	3.04	0.002	1.322

Note: Only age and log-transformed noun lemma frequency significantly contributed to the logistic regression model of naming accuracy.

model was significantly more predictive of naming accuracy than the Random model ($p < 0.001$), and the Main Effects model was more predictive than the Age model ($p = 0.007$). However, adding the age interactions (the Full model) did not significantly increase the model's predictive power ($p = 0.189$). This is likely due to near-ceiling effects, especially among the younger participants. Estimates of the Main Effects model, along with associated *p*-values and odds ratios, are shown in Table 3. Of the fixed variables, only age ($p < 0.001$) and target lemma frequency ($p = 0.002$) contributed significantly (see Figure 3). The odds ratio for age is about 98%, i.e. a decline in the likelihood of accurate responses of about 2% per year when other variables are held constant. To use more meaningful units, with each advancing decade, the odds of correct responding decline to 79% of the odds the decade before (i.e. a decline of 21% per decade). At age 25, the odds of accuracy are more than 4 times what they are at age 85. With all lexical variables set to their mean value, the odds of an accurate response for a 25-year-old are 132:1 (over 99%); these drop to 32:1 (about 97%) for an 85-year-old. Note that these numbers do not represent the *amount* by which accuracy is reduced, but the *change in odds* of accuracy, given a relatively high default level of accuracy.

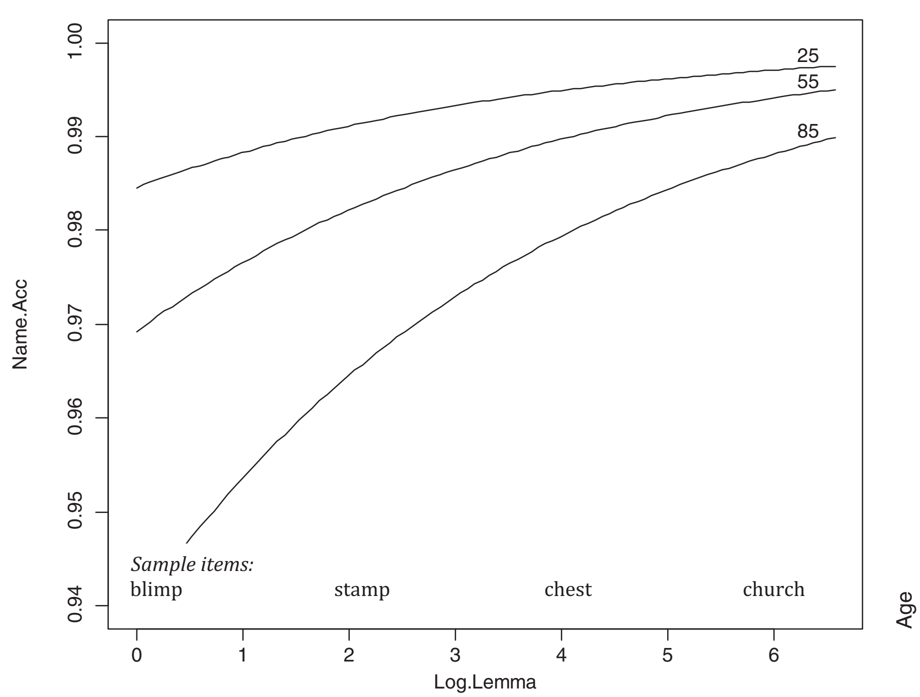

Figure 3. Full model predictions for naming reaction time by 25-, 55- and 85-year-olds.
Note: Only the main effect of Log Lemma Frequency was significant: Accuracy was higher for younger than older subjects and for more frequent than less frequent words.

This ceiling effect limits the strength of conclusions which can be drawn from the accuracy results. Nevertheless, the pattern found here supports expected effects of ageing and lexical frequency.

Latency of naming

Before conducting the regression analyses on reaction times, a final data clean-up step involved visually inspecting the distributions of RTs for each individual to identify outliers, defined as those which appeared to be unrepresentative of the word retrieval process for that individual. This approach is recommended by Baayen as a more valid way of detecting outliers than a single criterion value applied across individuals, and a way to avoid unnecessary data loss (Baayen, 2008). Across all participants, this process resulted in the exclusion of 33 additional RT data points.

Reaction times were transformed into *z*-scores to prevent age-related differences in average response speed and variability of response speed from giving rise to spurious interactions.[4] Mixed linear regression analyses showed that significant predictive power was gained by adding the Main Effects of the lexical variables to the Age model ($p < 0.001$), and by adding the age interactions (Main Effect model vs Full model; $p < 0.001$).

Table 4 shows the estimates of each variable included in the Full model. Reaction time is significantly influenced only by the interactions between Age and the following lexical variables: Age × Density Residuals ($p = 0.002$); and Age × Log Lemma Frequency ($p < 0.001$). Plots of each of these interactions in Figure 4 demonstrate that, consistent with previous research (e.g. Spieler & Balota, 2000), naming reaction times are shorter for more frequent items, and this effect becomes stronger with age.

The Density Residuals plot shows that the interaction with age is caused by a change in the direction of the effect from young to older adulthood. For young adults, ND residuals exert little influence. However, with age, increased ND appears to exert a more competitive effect, such that responses are slower for words with more neighbours than predicted by their length. The accuracy results, although not significant, were consistent with this change in direction; responses to higher-density items tended to be less accurate for the older participants. This illustrates that the latency findings are not due to a speed-accuracy trade-off. Once length and other variables are statistically factored out, words with more neighbours are retrieved more slowly, and perhaps less accurately, by older participants. The fact that frequency shows the expected facilitative effect, opposite to the obtained density effect and also increasing with age, provides additional confidence that the density effect is not attributable to shared variance with the other lexical variables.

Accounting for phonotactic probability

Although we took care to account, in various ways, for the lexical variables that have been shown to be confounded with ND, recall that our Density Residual measure was still significant related to mean phonotactic probability ($r = 0.494$, see Appendix 1). To factor out this relationship, we repeated the regression analysis, this time

Table 4. Model estimates for naming reaction time.

Naming reaction time (full model)				
Random effects	Variance	Std dev.		
Target	0.1815	0.4260		
Subject	0.0000	0.0000		
Residual	0.8088	0.8993		
Fixed effects	Coefficient	Std error	*t*-Value	*p*-Value
Intercept	−0.1288	0.2272	−0.567	0.5708
Age	−0.0030	0.0023	1.267	0.2051
Length (# Phonemes)	0.0983	0.0611	1.607	0.1081
Initial Phoneme Prob	−1.9770	1.3880	−1.424	0.1544
Mean Biphone Prob	−5.6170	13.5600	−0.414	0.6786
Density Residuals	−0.0049	0.0054	−0.909	0.3635
Frequency (Log.Lemma)	−0.0214	0.0271	−0.789	0.4304
Age*Phon	0.0003	0.0006	0.475	0.6347
Age*Initial.PP	−0.0189	0.0142	−1.338	0.1810
Age*Mean.BP	0.0575	0.1390	0.415	0.6782
Age*Dens.Resid	0.0002	<0.0001	3.097	0.0020
Age*Log.Lemma	−0.0009	0.0003	−3.175	0.0015

Note: The linear regression model of naming latency was significantly influenced by participant age and the interactions between age and two lexical variables: Density Residuals and Noun Lemma Frequency.

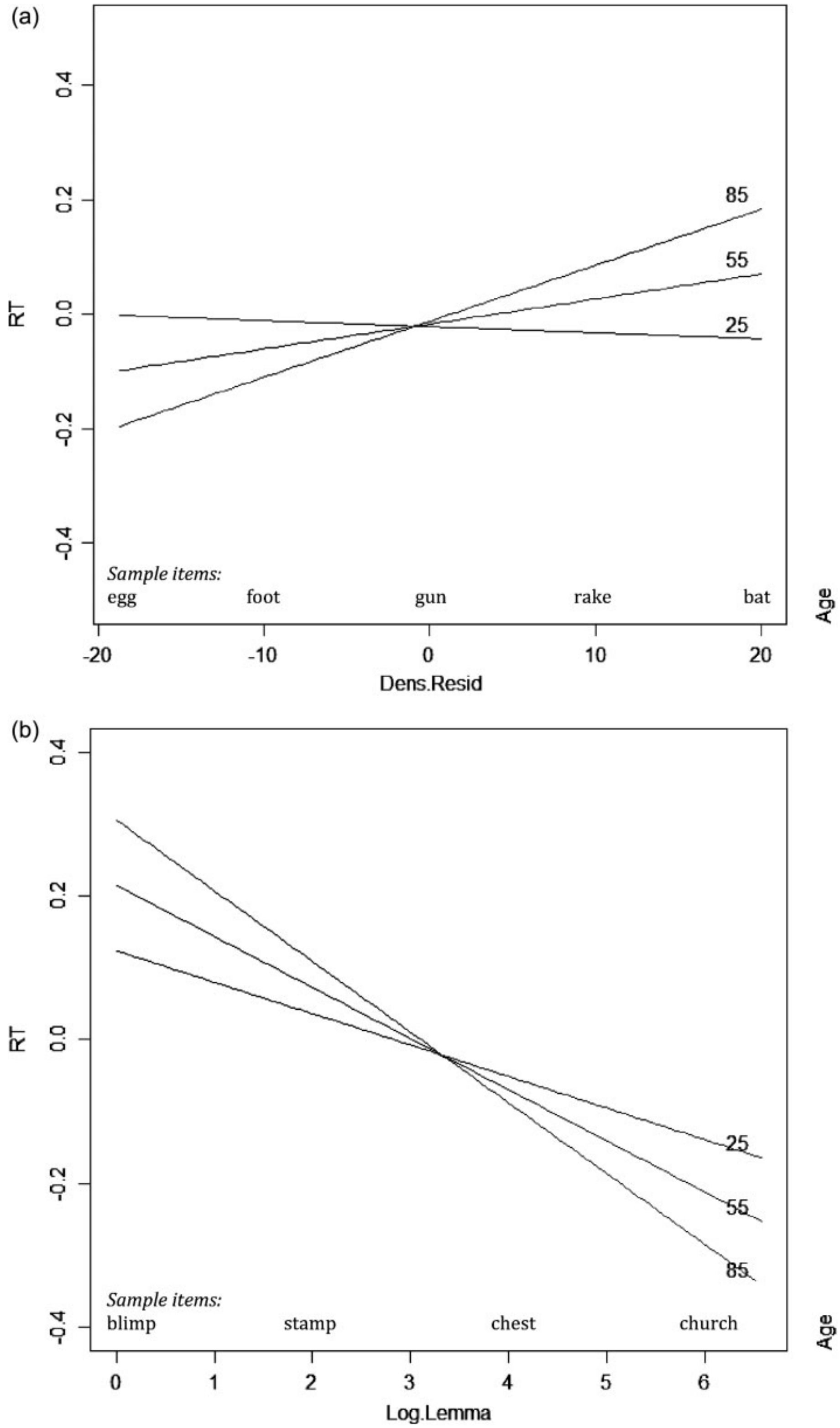

Figure 4. Full model predictions for naming reaction time by 25-, 55- and 85-year-olds.
Note: Age interacted with the lexical variables Density Residuals (a) and Log Lemma Frequency (b). Note that the regression analyses were run on *z*-transformed reaction times. Sample items are superimposed on each graph to illustrate the lexical variable on the *x*-axis. (a) Significant interaction between age and Density Residuals on Naming RT. (b) Significant interaction between age and Log Lemma Frequency on Naming RT.

regressing the Density Residuals (residual values after regressing Density on Length) first on Mean Phonotactic Probability. These "residuals of residuals" (called DensResid.PP) reflect the influence of ND beyond what would be predicted by length *or* by phonotactic probability. This density measure still retained a significant relationship to the original Density measure (r = 0.710), but was no longer significantly related to *any* of the other item variables: Length (r = −0.069); Initial PP (r = −0.069); Mean BP (r = −0.103); or Log Lemma Frequency (r = −0.006). Using this measure, we replicated the reaction time analyses. Although some of the p-values changed, the patterns of significance and non-significance were exactly the same in the RT models with DensResid.PP as in those with Dens.Resid. This provides further assurance that the pattern of results found represents the influence of ND, independent of other typically confounded item variables.

Exploring the lack of neighbourhood facilitation for younger subjects

The finding that ND residuals had very little effect on the younger subjects in our study was unexpected, and appears to contradict previous studies showing facilitative effects (e.g. A. H. D. Chan, K. Y. Chan, & Vitevitch, 2010; Vitevitch, 2002; Vitevitch et al., 2004). One difference might be that most of the adult studies have been conducted with undergraduates, almost always from introductory psychology courses. Thus, our youngest subjects were probably older than most of the participants in earlier studies. To explore this possibility, we used the Full model of naming latency generated from our results to predict the performance of "young" adults, i.e. 18-year-olds, the age of most freshman undergraduates. This involved taking the predictive values of the variables in the Full model, which were derived from our own naming latency data, and using them to generate patterns of naming latency performance for hypothetical subjects. Figure 5 suggests that facilitation at age 18 would be predicted by our results, but that this effect would weaken with age and disappear by 30 years of age (the mean age of the younger adult group), and gradually become competitive. Note that the scale of the y-axis in Figure 5 is different from the previous RT graphs, in order to highlight this contrast, so the absolute amount of change is quite small. What is important here is the demonstration that a change in direction of the effect is possible across adulthood, and is not inconsistent with other findings. In fact, it is not particularly surprising, given the instability of density findings in past studies. Thus, although the lack

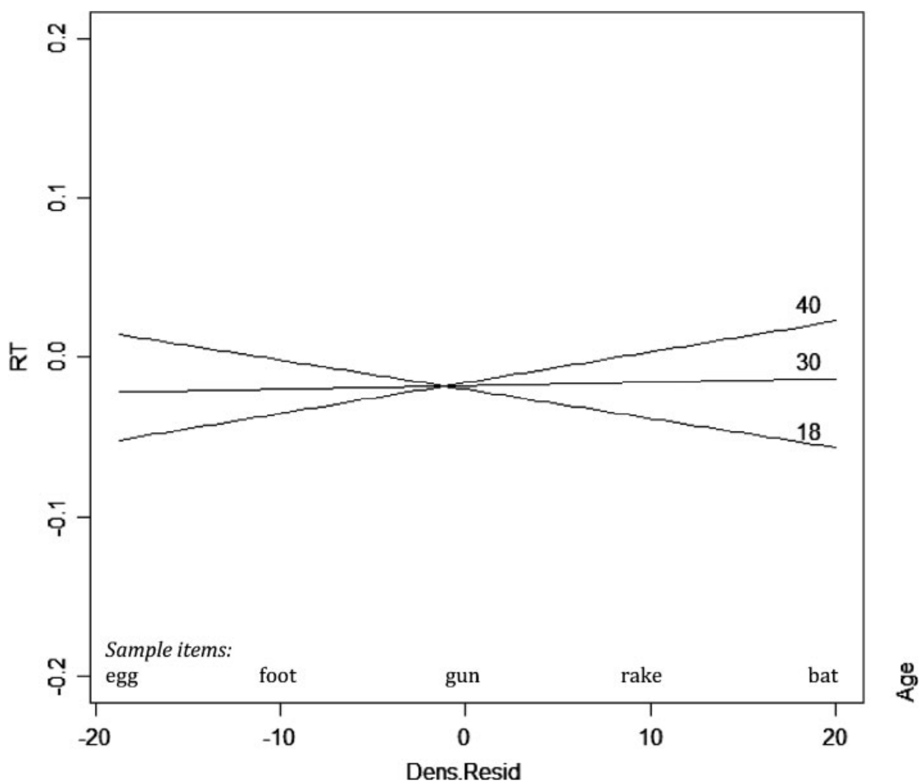

Figure 5. Model predictions for naming reaction time of younger-middle adults.
Note: Based on the current results, facilitative density effects may be predicted for "young young" adults (e.g. 18-year-old subjects), consistent with the results from previous studies. With age, however, facilitation is predicted to diminish and eventually reverse in direction.

of a density effect for our young adult subjects was not anticipated, our model predicts a pattern of performance consonant with existing results.

Exploring the neighbourhood competition in older subjects

Although previous studies have shown diminished facilitation by phonological neighbours for older adults, as described above, our findings are novel in illustrating that this may, under certain circumstances, result in phonological neighbourhood competition. One of the factors contributing to competition appears to be age. That is, the older the subjects, the weaker their lexical connections become, and the more likely it would be that neighbourhood effects would become competitive. In our study, we had 12 older participants in the "old-old" range (75–89 years) defined by Abrams, Trunk, and Merrill (2007). In a ToT elicitation study, they found that phonologically related words inhibited the resolution of ToT states for this age group (see Discussion). In our data, conducting a regression analysis on the naming latencies of just these oldest participants showed a close-to-significant inhibitory main effect of phonological ND (p = 0.059). Thus, just as with younger participant groups, the age range of subjects included in older groups may result in different patterns of findings, when the effect under examination varies continuously with age.

Summary and discussion

As illustrated by the correlational analyses, ageing clearly reduced both the speed and accuracy of word retrieval in this picture-naming task. When reaction times were normalised to factor out the main effects of age, regression analyses showed that age interacted with characteristics of the target, specifically its frequency and its phonological ND. The facilitative influence of frequency (here represented by the lemma frequency of the noun) showed the strongest influence on both outcome measures, consistent with previous studies showing a dominant effect of frequency (e.g. Dirks, Takayanagi, Moshfegh, Noffsinger, & Fausti, 2001; Kittredge et al., 2008; Taler et al., 2010). We also replicated the finding that the influence of frequency *may* increase with age (Spieler & Balota, 2000; but see Lagrone & Spieler, 2006; Newman & German, 2005 for differing results). Our results regarding ND in ageing were partially consistent with predictions from ISA models (Dell & Gordon, 2003), in showing a dissipation of the facilitative effect of neighbours with age, presumed to be due to a weakening of connections. This was shown in the interaction between age and ND residuals in determining naming response time, with the youngest participants showing no appreciable difference in response times to lower- and higher-density items, and the oldest participants showing considerably slower reaction times to higher-density items (see Figure 4).

Follow-up analyses showed that our data may still be consistent with studies showing facilitation for freshman-aged adults, although this facilitation would be predicted to disappear by about age 30. Interestingly, the one adult word production study by Vitevitch which found a competitive ND effect (Vitevitch & Stamer, 2006) recruited adult subjects from the community, leaving open the possibility that they may have been older (although no age range was provided). While the authors attributed the difference to the language tested (Spanish vs English), age differences should also be considered (and there may be additional differences between community based samples and undergraduate freshman samples that should be considered). Clearly, not all adults are alike, and response patterns may be influenced by sampling methods. This illustrates that it may not be valid to generalise from relatively homogeneous samples to the whole age range of adults.

A dissipation of phonological density facilitation throughout adulthood may complement prior results for younger adults, and is also consistent with previous predictions for older adults (Dell & Gordon, 2003; Gordon & Dell, 2001; Spieler & Balota, 2000; Vitevitch & Sommers, 2003). The mechanism for the change in density effects with age (i.e. weakening connection strength) has been previously proposed to account for alterations in lexical retrieval abilities, both in Burke's Transmission Deficit Hypothesis of lexical retrieval in ageing (Burke et al., 1991, 2004; Burke & Shafto, 2004), and with Dell and colleagues' (Dell et al., 1997; Dell & Gordon, 2003; Gordon & Dell, 2001) interactive activation (IA) model. In such models, the activation from shared phonemes of phonological neighbours spreads back to the target, strengthening its activation level and normally resulting in facilitation. However, when these connections are weakened, as is hypothesised to occur in ageing, the influence of phonological neighbours becomes increasingly competitive. When activation converges less effectively on the target, concurrently activated items are more likely to be produced instead of the target, predicting a reduction in accuracy (Dell & Gordon, 2003), and/or an increase in response latency. When activation converges less efficiently, it takes longer to resolve the activation of these competitors.

A hypothetical example based on the Dell and Gordon (2003) simulations can serve to illustrate this effect. Note that the inhibitory effect in the weight-lesioned network occurs at the lemma selection stage (see Table 1), reflecting interference from neighbouring activated words. Because activation reverberates throughout the network in an interactive fashion, these neighbours include both semantically related and phonologically related words. As noted above, the primary competitors in production tasks are semantic neighbours while, in the undamaged neighbourhood, phonological neighbours provide facilitative feedback to the target. In the dense network, then, the target word CAT might achieve an activation level of 100 units; its semantic neighbour, DOG, 60 units, and its phonological neighbours, CAN and MAT, 40 units each. In the empty network, there is no phonological neighbour to provide feedback to the target, so it ultimately receives less activation overall, say 80 units, while DOG continues to receive 60 units. If we assume that neighbours with activation levels within 30 units of the target cause interference (i.e. slowed retrieval of the target), then target retrieval will be slower in the empty network than in the dense network (neighbourhood facilitation). If, however, connection weights are weakened overall (e.g. everything is reduced by half its activation level), then in the dense network CAT would receive 50 units, DOG 30 units, and CAN and MAT 20 units each. In this case, all the neighbours are close enough to the target to cause interference at the lexical selection stage. In the empty network, CAT receives 40 units (still less than in the dense network, because of a lack of feedback from phonological neighbours), and DOG receives 30 units, resulting in only one competitor.

Note that no new processing mechanisms are hypothesised to account for these different outcomes. There are two opposing forces at play in all versions of this model: the existence of more neighbours provides more opportunities for activation to disperse away from the target. Under normal circumstances, however, the spread of activation ultimately accumulating on the target overcomes the competition from activated neighbours. But when the spread of activation is disrupted, activated neighbours are less able to provide facilitating feedback to the target. This dynamic trade-off between the cost and benefit of neighbours predicts a potential range of neighbourhood influence from facilitation to competition, depending on both the number of neighbours and the effectiveness of spreading activation within the lexical network.[5]

Consistent with these predictions, previous empirical findings of density effects in older adults have shown a reduction in the robustness of facilitation (Spieler & Balota, 2000; Vitevitch & Sommers, 2003). There is also some evidence to support the idea that co-activated phonologically neighbours may interfere in older speakers. Contrary to their expectations, Abrams and colleagues (Abrams et al., 2007) found differential effects of phonological relatedness on priming of ToTs for adults from different age groups. Specifically, for old-old subjects (75–89 years old), phonologically related primes from the same grammatical category as the target resolved fewer ToT episodes than did phonologically unrelated primes or phonologically related primes from a different grammatical category. For younger subjects, ToT resolution was facilitated by phonologically related primes, but only when they belonged to a different grammatical class than the target. Since all the targets were nouns in the current study, our findings are consistent with Abram and colleagues' finding of "an increased susceptibility in old-old adults for phonologically related words in the same grammatical class to become fiercer competitors" (p. 841). Furthermore, our account illustrates that it may not be necessary to propose specific damage to inhibitory mechanisms (such as lateral inhibitory connections) to account for such a result; it is predicted by IA dynamics, along with age-related weakening. (Of course, this does not mean that inhibition deficits might not *also* contribute to lexical retrieval deficits in ageing, and our results cannot rule out this possibility.)

Another major difference between the current study and previous studies which may have contributed to the different patterns of results is the number of stimuli included. This could have affected our results in two ways. First, the stimulus set used here is larger and, we propose, more representative of the lexicon (at least the single-syllable sector of the lexicon) than previous studies using small sets of carefully matched stimuli. Further research is required to determine whether our results for single-syllable words generalise to longer words. One challenge in addressing this question is to identify a meaningful measure of density, since most multisyllabic words have no neighbours as defined here. The second stimulus factor to consider is that requiring participants to name 400 items in a row may have exaggerated the impact of competitive influences among lexical items. Repeated naming may cause a cumulative build-up of activation, increasing local competition from the preceding test items (e.g. see Howard, Nickels, Coltheart, & Cole-Virtue, 2006; Oppenheim, Dell, & Schwartz, 2010). It remains to be seen whether such interference occurs for phonologically related as well as semantically related items in the ageing populations. We are investigating these possibilities in follow-up analyses.

A final factor distinguishing our study from previous work is the nature of the analysis. Unlike most previous ND studies, we treated our independent lexical variables as continuous rather than categorical factors (e.g. high density vs low density). When small sets of words (e.g. 10–15 per condition) are statistically compared, the power to detect differences in confounding variables is low. Thus, the target words in these studies may have been insufficiently controlled. We dealt with this problem by regressing density on its strongest confounds (length and phonotactic probability). This removed any variability due to these confounds, so was a more stringent test of density effects. Removing these sources of variability, which tend to have facilitative influences, may have allowed the competitive effects of density to emerge for the older participants.

Limitations

Despite our attempts to include a wide range of variability in our stimuli, our lexical variables and our subjects, we acknowledge a number of methodological limitations to the current study. There are many variables which have been shown to affect word retrieval, not all of which are explored here. In particular, semantic-level variables such as age-of-acquisition, familiarity and name agreement (e.g. Perez, 2007), and visual stimulus variables such as image complexity (e.g. Szekely & Bates, 2000) can influence naming accuracy or reaction time. We were primarily interested in lexical-phonological stages of word retrieval, and our name familiarisation task was intended to minimise the impact of these earlier stages on accuracy and reaction time. Nevertheless, they may still have influenced responses, and might help account for some of the variance not explained by our models.

It should also be acknowledged that most measures are somewhat crude approximations of the various factors influencing lexical retrieval. In particular, the use of frequency values gathered by Kuçera and Francis (KF, Francis & Kuçera, 1982; Kuçera & Francis, 1967) have come under criticism in recent years (e.g. Brysbaert & New, 2009), in part for being outdated. To assess whether the norms collected by these researchers (BN09) from video subtitles would be a better representation of lexical frequency, we correlated their word-form frequency measure with our noun lemma frequency measure derived from the KF norms. Although the raw values were not significantly related (r = 0.106, p = 0.134), log-transformed values were (r = 0.737, $p < 0.001$), suggesting that the log transformation compensates in large part for discrepancies in the different corpora. We also found that correlations between outcome measures (mean log RT and mean accuracy) and each frequency measure were not significantly different (ps > 0.70). Given this, we kept the KF measure in our analyses as a more relevant theoretical representation of lexical-semantic frequency; unlike BN09, our measure represented lemmas rather than word-form frequencies and counted only noun occurrences.

Another consideration is that the measure of ND used here accounts for the number of neighbours, but neglects the potential impact of certain other characteristics of phonological neighbourhoods, such as the "spread" of the phonological neighbourhood, i.e. the number of phoneme positions in a target that could be changed to form a neighbour (Vitevitch, 2007); neighbourhood clustering, i.e. the degree to which neighbours of a target are also neighbours of each other (A. H. D. Chan et al., 2010; K. Y. Chan & Vitevitch, 2009); and phonological distance (Mirman et al., 2010). Related to these variables is the strength of activation provided by the neighbours, based, for example, on the closeness of their relationship to the target. Chen and Mirman (2012) investigated this variable in a series of simulations. Interestingly, they showed that the direction of neighbourhood effects depends on strength of activation: strongly activated (close) neighbours exerted a net inhibitory effect and weakly activated (distant) neighbours a net facilitative effect. This principle accounted for different effects between near and far semantic neighbours (see also Mirman, 2011), between higher and lower frequency neighbours, between visual and auditory word recognition tasks, and between auditory word recognition and word production. The latter difference comes about because phonological neighbours are more strongly active during auditory word recognition tasks, in which they are directly activated from the acoustic input, than in word production tasks, in which they are activated indirectly, through feedback connections (Chen & Mirman, 2012).

The proposal that weakened connections result in competitive neighbourhood effects may seem to contradict the findings of Chen and Mirman (2012). However, as illustrated above in the hypothetical example of Dell and Gordon's network, it is not the *absolute level* but the *relative level* of target activation which is critical to its selection.[6] When the activation levels of the neighbours approach that of the target, neighbours become stronger competitors (Dell & Gordon, 2003). So when neighbours stop being facilitative, the mere fact of their activation creates competition. The lexicon's challenge then becomes resolving what Wheeldon calls the "riot of background activation" (2003, p. 82), that is, the co-activation of related items. Various mechanisms have been proposed to accomplish this, such as providing a jolt of activation to the most highly active unit (Dell et al., 1997), or building in lateral inhibitory connections at the lexical level, which are scaled by a sigmoid function to exaggerate the influence of more strongly active neighbours (Chen & Mirman, 2012). Such mechanisms suggest alternative parameters which may be manipulated to account for the interaction between age and density. Although the current results cannot distinguish among these possibilities, future simulations can address these questions.

Finally, we have considered only a small subset of a myriad of potential subject factors which might have influenced performance. In addition to age and education, more direct measures of cognitive function might reveal age effects that moderate naming performance. General cognitive effects of ageing, such as slowed processing and decreased inhibition, may underlie some language functions (Burke & Shafto, 2004), and specific language changes, such as increases in vocabulary (e.g. Verhaeghen, 2003), may moderate the effects of lexical characteristics on word retrieval.

Conclusion

By analysing naming responses to a large set of stimuli by adults ranging in age from 22 to 86 years, and by using mixed-model regression analyses which accommodate a number of different lexical variables and the full range of subject and item data, we provided a stringent test of the effects of phonological ND on word production. Results suggest that facilitative effects observed in previous studies may not be representative of the wider range of stimuli or subjects. In fact, as predicted by current IA models of lexical retrieval, neighbourhood facilitation effects were found to diminish with age, reversing direction as lexical connections weaken. It appears that, like all else, lexical neighbourhoods age, bringing about changes in the dynamics of interaction among the inhabitants.

Acknowledgements

This project was supported by Grant Number R03-DC007072 from the NIDCD of the National Institutes of Health. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIDCD or NIH. The authors also wish to acknowledge the valuable work of many students and research assistants, including Holly Kavalier, Megan Slater, Tracy Ball, Stephanie Leeper, Stephanie Cain, Zhen Chen, Ling-Yu Guo, Dawna Duff and Nichole Eden. Thanks also to Gary Dell, Amanda Owen, Karla McGregor and many anonymous reviewers for reading many earlier drafts of the paper.

Notes

1. Although Dell and Gordon (2003) simulated aphasia and we are examining normal ageing in the current study, we do not assume that word retrieval in ageing and aphasia are identical. However, we do think it likely, and a parsimonious hypothesis, that the mechanisms contributing to processing in both populations overlap.
2. Although previous studies have used summed probabilities, summing across phonemes and biphones exacerbates the extent to which these measures are confounded with length (see Storkel, 2004), so we used mean probabilities.
3. To ensure that manually measured and automatically measured RTs were similar, we compared manual and automatic RTs for a subset of 500 pseudo-randomly selected items. They showed a high positive correlation (r = 0.979), with an average discrepancy of 37 msec. In addition, paired t-tests across participants showed that mean naming response times were not significantly different with or without manual measurements for any of the 71 participants (*ps* from 0.278 to 1.0).
4. This possibility was suggested by an anonymous reviewer. As recommended by Wagenmakers, Krypotos, Criss, and Iverson (2012), we conducted the RT analyses using several scales (raw RTs, log-transformed RTs and z-score RTs). Results were the same, except that the interaction of Age and Initial Phoneme Probability was significant for RTs and log RTs, but not z-score RTs (and of course the main effects of Age and the Intercept were not significant for z-score RTs).
5. It is worth noting, as one of our anonymous reviewers pointed out, that this process of automatically spreading activation creates different dynamics within the lexicon than activation arising from external sources, such as through phonological priming. Although such studies can inform each other, there are many factors, such as the modality and timing of presentation, the goal of the task and features of the prime itself, that are likely to influence whether primes might facilitate or inhibit access to a target word (see Wheeldon, 2003). These are beyond the scope of the current paper, which deals only with internally generated spread of activation.
6. It is also important to keep in mind that what Chen and Mirman modelled was not a change in activation levels within a neighbourhood, but a comparison of neighbourhoods in which neighbours were relatively strong or weak compared to the target by virtue of the neighbourhoods' different structures. For example, a target from a tightly knit semantic neighbourhood would face competition, whereas a target with only distant semantic neighbours would benefit from facilitation. Because such structural characteristics are not expected to change with normal ageing, these relative effects should be maintained (i.e. closer semantic neighbours should still generate more competition than distant semantic neighbours). However, further simulations are required to determine what the net effect of weakened connections would be in such a model.

References

Abrams, L., Trunk, D. L., & Merrill, L. A. (2007). Why a superman cannot help a tsunami: Activation of grammatical class influences resolutions of young and older adults' tip-of-the-tongue states. *Psychology and Aging*, *22*, 835–845. doi:10.1037/0882-7974.22.4.835

Au, R., Joung, P., Nicholas, M., Obler, L. K., Kaas, R., & Albert, M. L. (1995). Naming ability across the adult life span. *Aging and Cognition*, *2*, 300–311. doi:10.1080/13825589508256605

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. New York, NY: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412. doi:10.1016/j.jml.2007.12.005

Baus, C., Costa, A., & Carreiras, M. (2008). Neighborhood density and frequency effects in speech production: A case for interactivity. *Language and Cognitive Processes*, *23*, 866–888. doi:10.1080/01690960801962372

Best, W. M. (1995). A reverse length effect in dysphasic naming: When elephant is easier than ant. *Cortex*, *31*, 637–652. doi:10.1016/S0010-9452(13)80017-2

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin*, *109*, 204–223. doi:10.1037/0033-2909.109.2.204

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*, 977–990. doi:10.3758/BRM.41.4.977

Burke, D. M., Locantore, J. K., Austin, A. A., & Chae, B. (2004). Cherry pit primes Brad Pitt: Homophone priming effects on young and older adults' production of proper names. *Psychological Science*, *15*(3), 164–170. doi:10.1111/j.0956-7976.2004.01503004.x

Burke, D. M., MacKay, D. G., Worthey, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word-finding failures in younger and older adults. *Journal of Memory and Language*, *30*, 542–579. doi:10.1016/0749-596X(91)90026-G

Burke, D. M., & Shafto, M. A. (2004). Aging and language production. *Current Directions in Psychological Science*, *13*(1), 21–24. doi:10.1111/j.0963-7214.2004.01301006.x

Chan, A. H. D., Chan, K. Y., & Vitevitch, M. S. (2010). Network structure influences speech production. *Cognitive Science*, *34*, 685–697. doi:10.1111/j.1551-6709.2010.01100.x

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 & Performance*, *35*, 1934–1949. doi:10.1037/a0016902

Chen, Q., & Mirman, D. (2012). Competition and cooperation among similar representations: Toward a unified account of facilitative and inhibitory effects of lexical neighbors. *Psychological Review*, *119*, 417–430. doi:10.1037/a0027175

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*, 335–359. doi:10.1016/S0022-5371(73)80014-3

Connor, L. T., Spiro, A., Obler, L. K., & Albert, M. L. (2004). Change in object naming ability during adulthood. *Journal of Gerontology: Psychological Sciences*, *59B*, P203–P209. doi:10.1093/geronb/59.5.P203

Dahlgren, D. J. (1998). Impact of knowledge and age on tip-of-the-tongue rates. *Experimental Aging Research*, *24*(2), 139–153. doi:10.1080/036107398244283

Dell, G. S. (1986). A spreading activation theory of retrieval in speech production. *Psychological Review*, *93*, 283–321. doi:10.1037/0033-295X.93.3.283

Dell, G. S., & Gordon, J. K. (2003). Neighbors in the lexicon: Friends or foes? In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (Vol. 6, pp. 9–37). Berlin: Mouton de Gruyter.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*, 801–838. doi:10.1037/0033-295X.104.4.801

Dirks, D. D., Takayanagi, S., Moshfegh, A., Noffsinger, P. D., & Fausti, S. A. (2001). Examination of the neighborhood activation theory in normal and hearing-impaired listeners. *Ear and Hearing*, *22*(1), 1–13. doi:10.1097/00003446-200102000-00001

Druks, J., & Masterson, J. (2000). *An object and action naming battery*. Philadelphia, PA: Taylor & Francis.

Evrard, M. (2002). Ageing and lexical access to common and proper names in picture naming. *Brain and Language*, *81*(1–3), 174–179. doi:10.1006/brln.2001.2515

Farrell, M. T., & Abrams, L. (2011). Tip-of-the-tongue states reveal age differences in the syllable frequency effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 277–285. doi:10.1037/a0021328

Feyereisen, P. (1997). A meta-analytic procedure shows an age-related decline in picture naming: Comments on Goulet, Ska, and Kahn (1994). *Journal of Speech, Language, and Hearing Research*, *40*, 1328–1333. Received from http://jslhr.asha.org/cgi/reprint/40/6/1328

Francis, W. N., & Kuçera, H. (1982). *Frequency analysis of English usage*. Boston, MA: Houghton Mifflin.

Frank, A. F., Tanenhaus, M. K., Aslin, R. N., & Salverda, A. P. (2007). *Frequency, neighborhood density, and phonological similarity effects in picture naming: An artificial lexicon study*. Paper presented at the 29th Annual Conference of the Cognitive Science Society, Austin, TX.

Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and Language*, *45*, 468–492. doi:10.1006/jmla.2000.2784

Goldrick, M., Folk, J. R., & Rapp, B. (2010). Mrs. Malaprop's neighborhood: Using word errors to reveal neighborhood structure. *Journal of Memory and Language*, *62*(2), 113–134. doi:10.1016/j.jml.2009.11.008

Gordon, J. K. (2002). Phonological neighborhood effects in aphasic speech errors: Spontaneous and structured contexts. *Brain and Language*, *82*(2), 113–145. doi:10.1016/S0093-934X(02)00001-9

Gordon, J. K., & Dell, G. S. (2001). Phonological neighborhood effects: Evidence from aphasia and connectionist modeling. *Brain and Language*, *79*(1), 21–23. doi:10.1006/brln.2001.2574

Harley, T. A., & Bown, H. E. (1998). What causes a tip-of-the-tongue state? Evidence for lexical neighbourhood effects in speech production. *British Journal of Psychology*, *89*(1), 151–174. doi:10.1111/j.2044-8295.1998.tb02677.x

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition*, *100*, 464–482. doi:10.1016/j.cognition.2005.02.006

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434–446. doi:10.1016/j.jml.2007.11.007

James, L. E., & Burke, D. M. (2000). Phonological priming effects on word retrieval and tip-of-the-tongue experiences in young and older adults. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *26*, 1378–1391. doi:10.1037//0278-7393.26.6.1378

James, W. (1890). *The principles of psychology*. Mineola, NY: Dover.

Kaplan, E., Goodglass, H., & Weintraub, B. (1978). *The Boston naming test*. Philadelphia, PA: Lea and Febiger.

Kittredge, A. K., Dell, G. S., Verkuilen, J., & Schwartz, M. S. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive Neuropsychology*, *25*, 463–492. doi:10.1080/02643290701674851

Kuçera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Lagrone, S., & Spieler, D. H. (2006). Lexical competition and encoding in young and older speakers. *Psychology and Aging*, *21*, 804–809. doi:10.1037/0882-7974.21.4.804

Landauer, T. K., & Streeter, L. A. (1973). Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. *Journal of Verbal Learning and Verbal Behavior*, *12*(2), 119–131. doi:10.1016/S0022-5371(73)80001-5

Luce, P. A., & Large, N. R. (2001). Phonotactics, density, and entropy in spoken word recognition. *Language and Cognitive Processes*, *16*, 565–581. doi:10.1080/01690960143000137

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear & Hearing*, *19*(1), 1–36. doi:10.1097/00003446-199802000-00001

Middleton, E. L., & Schwartz, M. F. (2010). Density pervades: An analysis of phonological neighborhood density effects in aphasic speakers with different types of naming impairment.

Cognitive Neuropsychology, *27*, 401–427. doi:10.1080/02643294.2011.570325

Mirman, D. (2011). Effects of near and distant semantic neighbors on word production. *Cognitive, affective & behavioral neuroscience*, *11*(1), 32–43. doi:10.3758/s13415-010-0009-7

Mirman, D., Kittredge, A. K., & Dell, G. S. (2010). *Effects of near and distant phonological neighbors on picture naming*. Paper presented at the Cognitive Science Society, Portland, OR.

Mortensen, L., Meyer, A. S., & Humphrey, G. W. (2006). Age-related effects on speech production: A review. *Language and Cognitive Processes*, *21*, 238–290. doi:10.1080/01690960444000278

Newman, R. S., & German, D. J. (2005). Life span effects of lexical factors on oral naming. *Language and Speech*, *48*(2), 123–156. doi:10.1177/00238309050480020101

Oppenheim, G. M., Dell, G. S., & Schwartz, M. S. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, *114*, 227–252. doi:10.1016/j.cognition.2009.09.007

Perez, M. (2007). Age of acquisition persists as the main factor in picture naming when cumulative word frequency and frequency trajectory are controlled. *Quarterly Journal of Experimental Psychology 60*(1), 32–42. doi:10.1080/17470210600577423

Roach, A., Schwartz, M. F., Martin, N., Grewal, R. S., & Brecher, A. (1996). The Philadelphia Naming Test: Scoring and rationale. *Clinical Aphasiology*, *24*, 121–133. Received from: http://aphasiology.pitt.edu/archive/00000215/

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object set: The role of surface detail in basic-level object recognition. *Perception*, *33*, 217–236. doi:10.1068/p5117

Sadat, J., Martin, C. D., Costa, A., & Alario, F.-X. (in press). Reconciling phonological neighborhood effects in speech production through single trial analysis.

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

Schneider, W., Eschman, A., & Zuccolotto, A. (2002b). *E-Prime reference guide*. Pittsburgh, PA: Psychology Software Tools.

Spieler, D. H., & Balota, D. A. (2000). Factors influencing word naming in younger and older adults. *Psychology and Aging*, *15*, 225–231. doi:10.1037/0882-7974.15.2.225

Storkel, H. L. (2004). Methods for minimizing the confounding effects of word length in the analysis of phonotactic probability and neighborhood density. *Journal of Speech, Language, and Hearing Research*, *47*, 1454–1468. doi:10.1044/1092-4388(2004/108)

Szekely, A., & Bates, E. (2000). Objective visual complexity as a variable in studies of picture naming. *Center for Research in Language Newsletter*, *12*(2), 1–33. Received from: http://crl.ucsd.edu/newsletter/12-2/article.html

R Development Core Team, (2011). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Taler, V., Aaron, G. P., Steinmetz, L. G., & Pisoni, D. B. (2010). Lexical neighborhood density effects on spoken word recognition and production in healthy aging. *Journal of Gerontology: Psychological Sciences*, 65B, 551–560. doi:10.1093/geronb/gbq039

Verhaeghen, P. (2003). Aging and vocabulary scores: A meta-analysis. *Psychology and Aging*, *18*, 332–339. doi:10.1037/0882-7974.18.2.332

Vitevitch, M. S. (1997). The neighborhood characteristics of malapropisms. *Language and Speech*, *40*, 211–228.

Vitevitch, M. S. (2002). The influence of phonological similarity neighbors on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 735–747. doi:10.1037/0278-7393.28.4.735

Vitevitch, M. S. (2003). The influence of sublexical and lexical representations on the processing of spoken words in English. *Clinical Linguistics & Phonetics*, *17*, 487–499. doi:10.1080/0269920031000107541

Vitevitch, M. S. (2007). The spread of the phonological neighborhood influences spoken word recognition. *Memory & Cognition*, *35*(1), 166–175. doi:10.3758/BF03195952

Vitevitch, M. S., Armbruster, J., & Chu, S. (2004). Sublexical and lexical representations in speech production: Effects of phonotactic probability and onset density. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 514–529. doi:10.1037/0278-7393.30.2.514

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science*, *9*, 325–329. doi:10.1111/1467-9280.00064

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, *40*, 374–408. doi:10.1006/jmla.1998.2618

Vitevitch, M. S., & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, & Computers*, *36*, 481–487. doi:10.3758/BF03195594

Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and Language*, *68*, 306–311. doi:10.1006/brln.1999.2116

Vitevitch, M. S., & Rodriguez, E. (2005). Neighborhood density effects in spoken word recognition in Spanish. *Journal of Multilingual Communication Disorders*, *3*(1), 64–73. doi:10.1080/14769670400027332

Vitevitch, M. S., & Sommers, M. S. (2003). The facilitative influence of phonological similarity and neighborhood frequency in speech production in younger and older adults. *Memory & Cognition*, *31*, 491–504. doi:10.3758/BF03196091

Vitevitch, M. S., & Stamer, M. K. (2006). The curious case of competition in Spanish speech production. *Language and Cognitive Processes*, *21*, 760–770. doi:10.1080/01690960500287196

Vitevitch, M. S., & Stamer, M. K. (2009). *The influence of neighborhood density (and neighborhood frequency) in Spanish speech production: A follow-up report*. Spoken Language Laboratory Technical Report. University of Kansas.

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. doi:10.1177/0023830908099070

Wagenmakers, E.-J., Krypotos, A.-M., Criss, A. H., & Iverson, G. (2012). On the interpretation of removable interactions: A survey of the field 33 years after Loftus. *Memory & Cognition*, *40*(2), 145–160. doi:10.3758/s13421-011-0158-0

Wheeldon, L. (2003). Inhibitory form priming of spoken word production. *Language and Cognitive Processes*, *18*(1), 81–109. doi:10.1080/01690960143000470

Appendix 1. List of 200 single-syllable items.

ant	boot	cheese	door	gate	key	nose	ring	snake	top
arm	bow	chest	dress	ghost	king	nun	road	sock	train
axe	bowl	church	drill	girl	kite	nurse	roof	sphinx	tree
bag	box	clock	drum	glass	knife	owl	rope	spoon	truck
ball	boy	clown	duck	globe	knot	pear	rose	stamp	vase
barn	bra	comb	ear	glove	lamp	pen	rug	stool	vest
bat	brain	corn	egg	goat	leaf	pie	saw	suit	vice
bear	bread	cow	eye	gun	leg	pig	scale	sun	watch
beard	bride	crab	fan	hair	lock	pill	scarf	swan	weight
bed	bridge	crib	fence	hand	log	pipe	screw	swing	well
bee	broom	cross	fire	harp	map	plant	scroll	tack	whale
bell	brush	crown	fish	hat	mask	plate	seal	tail	wheel
belt	bus	crutch	flag	hinge	match	plug	shark	tank	whisk
bench	cake	cup	floor	hoe	moon	pond	sheep	tape	wig
bird	can	dart	fly	hoof	moose	pool	shirt	tear	witch
blimp	cane	deer	foot	horse	mop	priest	shoe	tent	wolf
boat	car	desk	fork	hose	mouse	purse	sieve	thumb	worm
bomb	cat	dice	fox	house	nail	rain	sink	tie	wreath
bone	chain	dog	frog	jar	nest	rake	skunk	tire	wrench
book	chair	doll	fruit	jug	net	rat	snail	tongue	yak

Appendix 2. Results of correlational analyses.

a. Inter-correlations among lexical variables.

	Initial PP	Mean PP	Mean BP	Density B	Density Residual	Log Mean FreqB	Log Noun LemFreq	Log Noun LexFreq	Log All LexFreq
Phon	0.272*	0.121	0.355*	−0.643**	0.000	−0.636**	−0.231*	−0.183	−0.306*
InitialPP		0.480*	0.117	−0.039	0.178	−0.267*	−0.138	−0.107	−0.126
MeanPP			0.616**	0.301*	0.494*	0.090	−0.068	−0.032	−0.016
MeanBP				−0.063	0.205	−0.017	−0.043	−0.012	−0.008
DenB					0.767**	0.567**	0.119	0.114	0.264*
Density Residual						0.206	−0.039	−0.004	0.088
Log Mean FreqB							0.248*	0.216	0.313*
Log Noun LemFreq								0.964**	0.775**
Log Noun LexFreq									0.789**
Log All LexFreq									

b. Correlations of stimulus characteristics with raw outcome measures.

	Accuracy				Reaction time			
	All subjects	Young	Young-old	Older-old	All subjects	Young	Young-old	Older-old
Phon	−0.084	−0.001	−0.073	−0.146	0.149	0.128	0.146	0.138
InitialPP	0.012	0.021	0.033	−0.066	−0.101	−0.125	−0.020	−0.128
MeanPP	−0.004	−0.003	0.038	−0.047	−0.002	−0.078	0.053	0.011
MeanBP	0.001	0.041	0.046	−0.084	0.041	0.012	0.058	0.044
DenB	0.050	−0.032	0.039	0.098	−0.066	−0.105	−0.056	−0.041
DensResid	−0.006	−0.043	−0.011	0.005	0.041	−0.028	0.052	0.064
Log Mean FreqB	0.188*	0.099	0.158	0.230*	−0.248*	−0.253*	−0.218*	−0.207*
Log Noun LemFreq	0.217*	0.145	0.106	0.243*	−0.244*	−0.173	−0.189*	−0.295*
Log Noun LexFreq	0.214*	0.121	0.118	0.240*	−0.238*	−0.154	−0.192*	−0.298*
Log All LexFreq	0.228*	0.122	0.157	0.240*	−0.216*	−0.139	−0.187*	−0.265*

Notes: Phon = length in no. phonemes; Initial PP = phonotactic probability of initial phoneme; Mean PP = mean phonotactic probability of all phonemes; Mean BP = mean phonotactic probability of all biphones; DenB = phonological neighbourhood density; Density Residual = residual of DenB regressed on Phon; Log Mean FreqB = log-transformed mean frequency of neighbours; Log Noun LemFreq = log-transformed frequency of the noun lemma; Log Noun LexFreq = log-transformed frequency of the noun lexeme; Log All LexFreq = log-transformed frequency of all forms of the lexeme (see text for further explanation).

*Correlations significantly different from zero (criterion $r = 0.182$; $p < 0.01$, non-dir). **High correlations ($r > 0.500$).

Appendix 3. Density histograms showing distributions of each retained lexical variable for the 200 single-syllable items analysed.