Phonological Encoding in Sentence Production

Caitlin Hilliard (chillia2@u.rochester.edu), Katrina Furth (kfurth@bcs.rochester.edu),
T. Florian Jaeger (fjaeger@bcs.rochester.edu)
Department of Brain and Cognitive Sciences, Meliora Hall
Rochester, NY 14627 USA

Abstract

Previous tests of the phonological competition model (Dell, 1986) have mostly investigated the effects of phonological overlap (e.g. *pick-pin*) in isolated word production (e.g. primed picture naming). This is problematic since recent findings suggest that the effect of phonological overlap depends on the syntactic category of the phonologically related words, and few previous studies investigate phonological planning in the context of grammatical strings. We introduce a novel paradigm to examine two predictions of the so called parallel-then-sequential competition model (O'Seaghdha and Marin, 2000) against data from the distribution of disfluencies in sentence production. We also extend previous work by comparing different forms of phonological overlap (identity vs. similarity) in both word onsets and rhymes.

Keywords: phonological encoding; sentence production

Introduction

In order to speak, it is necessary to retrieve the segments that comprise each word (i.e. the phonemes) and organize them in the intended sequential order. This process, which is commonly referred to as phonological encoding, has been shown to be sensitive to interference from recently processed phonological material. A long line of research has employed these interference effects to infer the architecture of the system underlying phonological encoding. One of the most promising models that has emerged from this research is the phonological competition model and its offspring, which have been successfully applied to the distribution of speech errors, lexical production latencies, speech rate variations, and the tip-of-the-tongue phenomenon (Dell, 1986; Peterson, Dell & O'Seaghdha, 1989; Sevald and Dell, 1994; O'Seaghdha and Marin, 2000).

The original phonological competition model assumed that all segments are activated in parallel (Peterson, Dell & O'Seaghdha, 1989). The model contains a network comprised of nodes for both words and phonemes with top-down and bottom-up feedback connections between the nodes. During planning of a spoken word, activation spreads downward from the word node to the phonemes comprising the word. Activation also travels upwards through feedback connections from phonemes to the word nodes, leading to activation of word nodes phonologically related to the target. It is not until all the segments of a word have been assigned to their appropriate position in the syllable frame that the speaker starts articulating the word.

Evidence from later work suggests a sequential component is required. This evidence comes from studies on the effect of phonological overlap between adjacent words (Sevald & Dell, 1994; O'Seaghdha & Marin, 2000). Participants were asked to produce sequences of mono-syllabic consonant-vowel-consonant words as many times as possible within an 8 second period, where the adjacent words contained identical onsets (e.g. PICK-PIN) or identical rhymes (e.g. PICK-TICK). While onset overlap strongly inhibited (slowed down) production, rhyme overlap led to less inhibition (O'Seaghdha and Marin, 2000) or even facilitation (Sevald & Dell, 1994). To account for these findings, O'Seaghdha and Marin (2000:59) propose the parallel-then-sequential competition model, according to which "all segments are first activated, providing feedback to form-related words, and then selected in sequence." This would account for inhibition found in end-related conditions: parallel activation of the phonemes of a word retrieved for articulation spreads upwards through feedback connections to words that share these phonemes in the same position. If one of these words has recently been produced, its relatively high lingering activation combined with the feedback activation slows down selection of the target word. This effect is larger for overlapping onsets than rhymes because more time passes before the rhymes are required (that is the sequential component of the model).

Another prediction of the phonological competition model is that effects of phonological competition are dependent on the frequency of the target. High-frequency words are thought to encounter greater competition due to the rapid activation of their segments, which in turn is assumed to cause more competition between discrepant syllables from adjacent words. This prediction has received support from a variety of paradigms, including naming and word pair production (O'Seaghdha & Marin, 2000).

However, previous tests of the phonological competition model have almost completely been limited to isolated word production. The interaction between frequency and phonological overlap, which is crucial for the parallel-then-sequential model, has so far only been tested in isolated word production. This is problematic since recent evidence suggests that the effect of phonological overlap depends on both the syntactic category of the phonologically related words and the order, which was not controlled in the works cited above. For example, Janssen and Caramazza (2009) find inhibition due to onset overlap only for noun-noun sequences, and facilitatory effects in adjective-noun, noun-verb, and adjective-adjective-noun sequences. They hypothesize that grammatically common sequences (noun-verb, adjective-noun, etc.) show different effects than sequences that are rare (e.g. noun-noun). This result would call into question the interpretation of previous evidence in support of the parallel-then-sequential competition model.

We present two production experiments that seek to test the predictions of the parallel-then-sequential competition model against data from phonological encoding in complete sentences. Experiment 1 investigates the effect of onset overlap; Experiment 2 rhyme overlap. In both experiments we test the prediction that more frequent target words should be more severely inhibited than less frequent target words. The paradigm we introduce was designed to allow us to investigate phonological encoding in the context of sentence production.

Experiment 1: Onset Overlap

Participants were asked to produce simple transitive sentences. We manipulated the phonological overlap between the verb (lost, below) and the object noun. The object either (1a) had the same initial consonant and vowel as the verb (lost-lock), (1b) had the same vowel, but the initial consonant differed in exactly one phonological feature (lost-rock), or (1c) had the same vowel and the initial consonant did differ in at least two phonological features (lost-sock).

(1) (a) The maid **lo**st the **lo**ck [identical onset]
(b) The maid **lo**st the **ro**ck [similar onset]
(c) The maid **lo**st the **so**ck [unrelated control]

Our design hence differs in several potentially relevant aspects from previous work. First, previous work has generally compared phonological overlap between words of the same syntactic category (some recent exceptions are Abrams & Rodriguez, 2005; Damian & Dumay, 2007; Janssen & Caramazza, 2009). Second, previous work has almost exclusively focused on the comparison between onset identity and phonologically unrelated words. Little is known about the role of high phonological similarity in the absence of identity (but see Roelofs, 1999) –despite strong typological evidence that there is a bias against adjacent similar, but not necessarily against adjacent identical units (for summary, see Graff, 2010). Here, we compare identical to similar to unrelated words. Third, to address the possibility that phonological context plays a role in how phonological relatedness between proximate words affects production, we keep the preceding word and the rhyme of the target word identical across the conditions of an item (_ock).

Method

Participants Thirty University of Rochester undergraduates participated in the experiment for $10 compensation each. All reported to be native speakers of English.

Materials The experiment consisted of 36 items and 54 fillers. Items were simple transitive sentences in the three conditions illustrated in example (1) above. Participants saw one of three Latin-square-designed lists, so that each participant saw each item exactly once and equally many instances of each condition.

Item triplets were created using the MRC Psycholinguistic database (Coltheart, 1981). We first extracted all mono-syllabic noun triplets that differed in only in their phonological onset. For each triplet we searched for a transitive verb, so that (a) the verb's first syllable overlapped with one of the nouns in terms of its initial consonant and nucleus, but not in terms of its final consonant (identity condition) and (b) the verb's onset differed from the onset of one of the other nouns in only one phonological feature (similarity condition), and (c) the last noun's onset differed from the verb's onset in at least two phonological features. The resulting verb-noun-triplets were filtered to remove combinations with very infrequent nouns or verbs. Of the remaining materials, we chose those 36 combinations that we deemed best suited in terms of the intuitive plausibility of the verb object combination (e.g. … lost the lock/rock/sock. in (1) above).

We then constructed simple transitive sentences based on these verb-noun-triplets. The constraints of the design forced us to reuse five noun triplets with 11 different verbs (e.g. "The traveler bashed her sack/pack/ back" and "The guest padded his sack/back/pack"). However, stimuli were distributed in such a way that participants never actually saw the same noun twice in the experiment.

Log-transformed frequency information for each verb and object was obtained from the 400 million word Contemporary Corpus of American English (Davies, 2008). Additionally, we gathered information about several variables known to affect lexical production for all item triplets to avoid potential confounds. The following control properties of both the verb and object noun were considered: the log-transformed frequency of the first syllable across all words of English both in first position only and in any position (obtained from the CELEX lexical database, Baayen, Piepenbrock & Gulikers, 1995), the neighborhood density1, and frequency-weighted neighborhood density (Vaden, Hickok & Halpin, 2009). Controls were balanced across conditions. We also conducted mixed effect regression analyses (not reported here due to limited space) that confirmed that none of the results reported below were caused by main effects of the control variables. Additionally, we gathered plausibility norms for each sentence over Amazon's online service Mechanical Turk.

For the target stimuli, subjects were always third person lexical nouns. Half of the verbs occurred in the present tense and half in the past tense (e.g. lost in (1) above). Twenty-seven (75%) of the verbs were monosyllabic, and the other 9 verbs were disyllabic. Object nouns were always monosyllabic. The coda (e.g. /k/ in (1) above) was held identical across the three conditions.

Since previous work suggests that phonological relatedness affects production most strongly (or, possibly, only) when accompanied by orthographical overlap (Damian & Bowers, 2003), we tried to accompany

[1] See http://128.252.27.56/Neighborhood/NeighborHome.asp.

phonological overlap by orthographic overlap as much as possible within the limits defined by the need to create a sufficiently large set of plausible stimuli. The orthography of verb-object onset consonant in the identity condition (‘l’ in (1) above) was always matched. The orthography of the vowel was matched between verb-object pairs in 28 of the 36 items (if it was matched for an item it was matched for all conditions of that item). The orthography of the coda was matched in all but two items, for which some conditions contained a silent “b” (as in limb and lamb).

Procedure Participants were instructed that they would have to solve math problems while remembering sentences. There were two minimally different variants of the task. Half of the participants saw variant A, the other half saw variant B. Since the results did not differ between the two versions of the task (there were no interactions between the version of the experiment and the results reported below), we pool the data from the two versions.

In variant A, trials consisted of three phases. In the formation phase, participants saw the subject noun phrase (e.g. the maid), the verb (e.g. lost), and the object noun phrase (e.g.) the lock arrange in a triangle, randomly positioned from trial to trial. Participants formed a sentence from these parts and spoke it aloud. There was no time pressure. In the distraction phase, participants read two math problems aloud and typed the answer in a box on the screen. In the recall phase, one word from the original sentence was presented, prompting recall of the entire sentence as accurately as could be remembered. For target trials the prompt was always the subject of the sentence.

The procedure of variant B was identical to that described above except for two aspects introduced to increase the overall rate of disfluencies. First, we introduced time pressure during both formation and recall, with a green progress bar indicating remaining time. Based on piloting, the deadlines were set at 5500 milliseconds and 1750 milliseconds, respectively. To further increase the task's complexity, a temporal adverbial (e.g. yesterday) was added to the formation phase (the four words were now positioned in a diamond). In line with our expectations, version B resulted in more overall disfluencies.

The experiment began with a short practice session (12 trials) after which participants were instructed to ask the experimenter any questions they had about the procedure. Practice trials only contained phonologically unrelated stimuli. The practice session was followed by the experimental session (90 trials).

Scoring Productions from both the formulation and the recall phase of all item trials were transcribed and annotated by an undergraduate RA. Annotation reliability was ascertained by comparing annotation of 100 sentences against those of the first author. Since the disfluency rate on recall trials was very low (5%), we report only the analysis of formation trials. The presence of disfluencies was recorded for four separate regions: (1) before the subject, (2) after the subject and before the verb, (3) after the verb and before the head noun of the direct object, and (4) in the remainder of the sentence. This is illustrated in example (2) by squared brackets, where there is one disfluency in the subject region (a restart), no disfluency in the verb region, two disfluencies (one filled pause, one word lengthening) in the object region, and no disfluency in the remainder of the sentence.

(2)	Subject	Verb	Object	Remainder
Example	[*The, the maid*]	[*lost*]	[*the, uh, lo-ock*]	[*yesterday*]
Disfluency	yes	no	yes	no

The data from two participants had to be excluded because they had to abort the experiment or did not follow the instructions. Incomplete trials were excluded from the analysis. All data from one item had to be excluded from the analysis because of a copy and paste error in the experimental lists. This left 994 formation trials with an average disfluency rate across all regions of 16%.

Analysis

To analyze the distribution of disfluencies by conditions, we conducted separate mixed logit regression analyses (Jaeger, 2008) for each sentence region for both the formation and the recall trials. In all cases, a random intercept for participants was the only justified random effect. We first fit a model with the (a) the random effect, (b) a covariate for plausibility, (c) the three-way design factor, (d) verb and object frequency and (e) the interaction between (c) and (d). The plausibility control was included to avoid potential confounds since its effect reached significance in some analyses. We then assessed the significance of variables by comparing the full model against a model without that variable. Directions of effects were assessed by means of planned comparisons (e.g. we always used Helmert coding for the design factor, comparing the Identical < Similar < Unrelated condition).

Results

In the subject region, there were no significant effects. *In the verb region*, two main effects and an interaction were observed. Unsurprisingly, participants were less likely to produce a disfluency the higher the logarithm transformed verb frequency ($\beta=-.14$, $z=-2.1$, $p<.04$), and the higher the sentence’s plausibility rating ($\beta=-.23$, $z=-2.2$, $p<.03$). There was no main effect of phonological overlap between the verb and object. There was, however, an interaction between phonological overlap and verb frequency ($\chi^2(3)=11.1$, $p<.02$). Planned comparison revealed that, with increasing verb frequency, both similar and identical verb-object onsets led to increased probability of a disfluency compared to the control condition ($\beta=.36$, $z=2.5$, $p<.02$). The effect of verb frequency did not differ between the two overlap conditions ($p>.9$).

In the object region, a main effect of phonological overlap was observed ($\chi^2(3)=6.4$, $p<.05$). This effect was driven by the identical verb-object onsets, which were associated with a higher rate of disfluencies ($\beta=.66$, $z=2.4$, $p<.02$) compared to both similar verb-object onsets and the control condition (the latter two conditions did not differ, p>.4). Additionally, there was a significant interaction of phonological overlap with object frequency ($\chi^2(3)=10.8$, p<.01). This interaction went in the opposite direction of the effect observed over the verb region: the effect of identical onsets decreased with increasing object frequency (β=-.38, z=-2.6, p<.01). A marginal interaction of verb frequency with similar onsets in the same direction was also observed (β=-.29, z=-1.8, p<.07).

Discussion

In comparing the effects of Experiment 1 to previous work, it is important to keep in mind that the paradigm employed here likely has considerably less power to detect effects than studies that examine adjacent word produced in isolation (rather than using a comprehended prime).

The inhibitory results found in the verb region between related onsets are in line with previous work (O'Seaghdha & Marin, 2000; Sevald & Dell, 1994) and provide further support for the phonological competition model. These studies (Sevald & Dell, 1994; O'Seaghdha & Marin 2000) had participants repeat phonologically related words in fast succession as many times as possible. Since our task did not involve so much phonological repetition, it is unsurprising that the hypothesized effect of phonological relatedness may not reach significance in all cases.

However, the interaction between phonological relatedness and object frequency for the object region exhibited the opposite of the predicted pattern. We consider two interpretations of this effect. First, it is possible that the phonological inhibition model needs to be revised. Several other studies have found that the effects of phonological overlap depend on whether the two phonologically related words are of the same syntactic category (Abrams & Rodriguez, 2005; Damian & Dumay, 2007; Janssen & Caramazza, 2009). Most of this work has employed paradigms in which one of the two related words (the so called prime) is only comprehended. This differs from the procedure employed here, where both words (the verb and the object) are produced.

It is also possible that the different interactions of frequency and phonological relatedness are due to properties of the procedure used here. We return to this point below.

With regard to the relation between similarity and identity, the results are relatively weak but consistent. With regard to the interaction with verb frequency (in the verb region), similar and identical onsets group together compared to the control condition. With regard to the interaction with object frequency (in the object region), a similar pattern was observed, although the effect here was stronger for identical than for similar onsets. The same ordering is found for the main effect of phonological overlap at the object. Taken together, these results suggests that at least at some stage, subphonemic overlap affects lexical production and that partial feature overlap can create effects similar to complete feature overlap (i.e. phoneme identity).

Experiment 2: Rhyme Overlap

Experiment 2 follows the design of Experiment 1, but investigates the effects of phonological overlap between rhymes. An example item is shown as (3) below. As spelled out in the introduction, the parallel-then-sequential competition model predicts less inhibition for related rhymes than for related onsets.

(3) (a) The owners scr**ub** the p**ub** [identical rhyme]
(b) The owners scr**ub** the p**up** [similar rhyme]
(c) The owners scr**ub** the p**us** [unrelated control]

Method

Participants Thirty undergraduate students from the same population as in Experiment 1 participated for $10 compensation each. All reported to be native speakers of English.

Materials Following Experiment 1, there were 36 items and 54 fillers. Item triplets were created to mirror those from Experiment 1, differing only in the location of overlap (the rhyme instead of the onset). The frequency of the verb and object, as well as all the controls variables mentioned in the description of Experiment 1 were elicited for Experiment 2.

For target stimuli, subjects were again always third person lexical nouns. Thirty-one of 36 verbs occurred in the present tense, due to less likelihood of participants altering the coda of either the verb or object for sentences with present tense verbs. Half of verbs were monosyllabic and half disyllabic. Object nouns were always monosyllabic. The onsets of the nouns within each triplet were held consistent. The orthography of verb-object coda consonant was always matched. The orthography of the vowel was matched between verb-object pairs in 26 of the 36 items. The orthography of the onset within triplets was matched in all items.

Procedure The procedure was identical to that of Experiment 1, having two minimally different variants A and B, with the same differences as in Experiment 1. Again the data from each variant were pooled, as the results of interest do not differ between the two versions.

Scoring Scoring was identical to Experiment 1, with one difference: trials in which participants introduced additional material in the coda of the verb or object (e.g. *scrubbed* for *scrub, pubs* for *pub)* were excluded from the analysis. Data from four items had to be excluded from the analysis since less than 40% of the participants managed to construct the intended sentence. This left 656 formation with an average disfluency rate across all regions of 21%.

Results

In the subject region, two main effects and an interaction were observed. Parallel to Experiment 1, participants were less likely to produce a disfluency the higher the sentence's plausibility rating (β=-.25, z=-1.9, p<.06). There also was a marginal main effect of phonological overlap ($\chi^2(3)=4.6$, p=.09), so that both identical and similar rhymes were more likely to lead to a disfluency compared to the control condition (β=.57, z=2.2, p<.03). The two phonological overlap conditions did not differ from each other (p>.9). Additionally, there was a marginal interaction between phonological overlap and verb frequency ($\chi^2(3)=4.9$, p=.09) for identical compared to similar rhymes and the control condition, participants were more likely to produce a disfluency the higher the verb frequency (β=.36, z=2.2, p<.03). The effect of verb frequency did not interact with the difference between similar rhymes and the control condition (p>.9).

In the verb region, there were no effects. *In the object region*, there was no main effect of phonological overlap or object frequency, but the two-way interaction between phonological overlap and object frequency approached significance ($\chi^2(3)=5.1$, p<.08): for identical compared to unrelated rhymes, participants were again more likely to produce a disfluency the higher the object frequency (β=.40, z=2.2, p<.03). The effect of object frequency did not differ between identical and similar onsets (p>.8).

Discussion

The results in the subject region of Experiment 2 are similar to the findings in the verb region in Experiment 1; conditions with both identical and similar rhymes increase the likelihood of producing a disfluency. Also in line with Experiment 1, the probability of producing a disfluency between identical rhymes increased with verb frequency. We can only speculate that the reason that these effects showed up in the subject region in Experiment 2, but verb region in Experiment 1, is due to the relatively low power our paradigm has compared to previous work (see above).

In the object region, disfluencies were more likely in the identical rhyme condition with increasing object frequency, which contrasts the result from Experiment 1.

General Discussion

The results obtained in Experiments 1 and 2 are mixed, although overall they tend to support the phonological competition theory. First, only inhibitory main effects of phonological overlap were observed. Second, the interaction of phonological overlap with frequency went in the direction predicted by the parallel-then-sequential competition model (O'Seaghdha and Marin, 2000) for three out of the four significant interactions.

With regard to the comparison of similar vs. identical onsets and rhymes, we found that identity is generally worse than similarity. Identical onset and rhymes lead to inhibition compared to similar ones. Words with similar onsets or rhymes mostly grouped with the control condition, except with regard to the interaction with frequency. This suggests competition may also take place at the subphonemic level.

Two findings were not predicted by the parallel-then-sequential competition model. First, we do not find that onsets lead to more inhibition than rhymes: The coefficients of the phonological overlap effects are comparable across the two experiments. Additionally, disfluency rates were higher overall in Experiment 2. Second, Experiment 1 returned one significant interaction of phonological relatedness and frequency in the *un*expected direction: for the object region, the inhibitory effect of onset overlap *de*creased with increasing frequency of the object

These unexpected effects could be due to a drawback of our paradigm to be addressed in future work: the lack of control over the time course of comprehension and subsequent production of the sentence. Figure 1 illustrates the two endpoints of the spectrum that describes the potential overlap of cascading comprehension and production processes in our paradigm. The hypothetical speaker in the left panel, initiates talking as soon as a likely candidate for the subject phrase (*maid*) has been comprehended. As a consequence, parts of the to-be-produced sentence are still being comprehended while articulation is already initiated. The hypothetical speaker in the right panel, initiates production planning only after all parts of the stimulus have been comprehended, thereby avoiding overlap between comprehension and production.

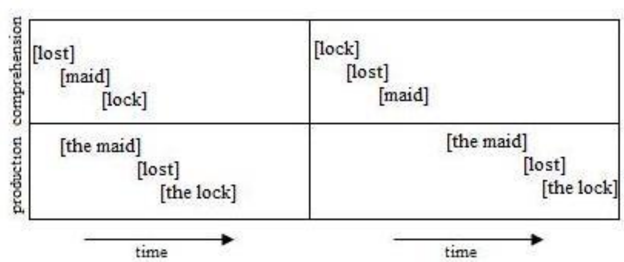

Figure 1: Illustration of different hypothetical time courses of stimulus comprehension and production in our paradigm

These different time courses are likely to lead to different results: phonological overlap has been shown to differ depending on whether verbal apprehension (in particular, of written input) co-occurs with production of the target word (O'Seaghdha & Marin, 2000 vs. Taylor & Burke, 2002), sometimes turning inhibitory into facilitatory effects. Since we did not control the strategies employed by participants, it is possible that the resulting mix of strategies weakened effects or even caused the unexpected interaction between phonological overlap and frequency in Experiment 1.

We must also factor in that the different syntactic categories of the two words could have produced some of the facilitatory effects. When the phonologically related prime and target are of different syntactic categories this facilitates TOT state resolution compared to the presentation of same-category primes, which has no effect (Abrams & Rodriguez, 2005; Burke, MacKay, Worthley & Wade, 1991). In another study examining these effects, phonological onset overlap yielded different effects based on word order; inhibitory effects are found in noun-noun utterances, while facilitatory effects are found in adjective-noun utterances (Janssen & Caramazza, 2009). Janssen and Caramazza link this difference to differences in the relative time course of grammatical and phonological encoding for grammatically typical (e.g. adjective-noun) and grammatically atypical (e.g. noun-noun) utterances. In our experiments, all utterances only contained grammatically typical orders. Yet, it is possible that the overlapping phonological encoding of the verb and object took place during different time points of their respective grammatical encoding, explain the results of Experiment 1. Without further stipulations, it would, however, not account for the observed inhibition for both the verb and the object when their rhymes overlapped (Experiment 2). In short, our data suggest a three-way interaction of phonological encoding, grammatical encoding, and they type of phonological overlap (onset vs. rhyme). We conclude that the account introduced by Janssen and Caramazza (2009) requires further modification.

Despite the drawbacks discussed above, we consider the paradigm introduced here a promising first step toward examining phonological competition effects during sentence production. While the lexical content to be produced by participants in our experiments was still largely scripted, participants had to create the syntactic structure themselves from the parts presented to them. Recall also that the adverbial in our stimuli (e.g. *yesterday*) could be produced in several positions of the sentence, highlighting the fact that grammatical encoding of the sentences was not fully scripted, unlike the majority of over 200 previous experiments we recently surveyed (but see Bock, 1987; Jaeger, Furth, & Hilliard, 2011).

In ongoing work, we are exploring ways to increase the statistical power of the approach by analyzing word duration and speech rate data from the different sentence regions. Word durations of the determiner before the object noun, for example, should be lengthened when lexical production of the noun is inhibited.

Acknowledgments

The authors are grateful to Andrew Watts for programming the experiments, to S. Brown, J. Ferris and R. Glass for running participants, and to S. Kumar, and R. Murphy for transcription and annotation. Experiment 1 was conducted by KF as part of her B.A. thesis, Experiment 2 was conducted by CH. This work was partially funded by a Bishop Award to KF and by a Alfred P. Sloan Research Fellowship and the Wilmot Award to TFJ.

References

Abrams, L., & Rodriguez, E. L. (2005). Syntactic class influences phonological priming of tip-of-the-tongue resolution. *Psych Bulletin & Review*, 12(6), 1018-1023.

Abrams, L., White, K. K., & Eitel, S. L. (2003). Isolating phonological components that increase tip-of-the-tongue resolution. *Memory & Cognition*, 31, 1153-1162.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX Lexical Database (Version Release 2) [CD-ROM]. Philadelphia: LDC, University of Pennsylvania.

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *JML*, 30(5), 542-579.

Coltheart M (1981) The MRC psycholinguistic database. *JEP*, 33A: 497-505.

Damian, M. F., & Bowers, J. S. (2003). Effects of orthography on speech production in a form-preparation paradigm. *JML*, 49(1), 119-132.

Damian, M.F., & Dumay, N. (2007). Time pressure and phonological advance planning in spoken production. *JML*, 57(2), 195-209.

Davies, M. (2008) The Corpus of Contemporary American English (COCA): 385 million words, 1990-present.

Dell, G. (1986). A spreading activation theory of re- trieval in language production. *PsychReview*, 93, 282321.

Graff, P. (2010). Towards a Feature-Based Model of Phonological Similarity. Cambridge, MA: MIT.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *JML*, 59(4), 434-446.

Jaeger, T. F., Furth, K., & Hilliard, C. (2011). *Phonological encoding during unscripted sentence production*. Unpublished manuscript, University of Rochester.

James, L. E., & Burke, D. M. (2000). Phonological priming effects on word retrieval and tip-of-the-tongue experiences in young and older adults. *JEP:LMC*, 26, 1378-1391.

Janssen, N., & Caramazza, A. (2009). Grammatical and Phonological Influences on Word Order. *Psych Science*, 20, 1262-1268.

O'Seaghdha, P. G., & Marin, J. W. (2000). Phonological competition and cooperation in from-related priming: sequential and nonsequential processes in word production. *JEP: HPP*, 26, 57–73.

Peterson, R.R. Dell, G.S., & O'Seaghdha, P.G. (1989). A connectionist model of form-related priming effects. *11th CogSci*.

Roelofs, A. (1999). Phonological segments and features as planning units in speech production. *LCP*, 14, 173-200

Sevald, C.A., & Dell, G.S. (1994). The sequential cuing effect in speech production. *Cognition*, 53, 91–127.

Starreveld, P. A. (2000). On the interpretation of onsets of auditory context effects in word production. *JML*, 42, 497–525.

Taylor, J.K., & Burke, D.M. (2002). Asymmetric aging effects on semantic and phonological processes: Naming in the picture-word interference task. *Psych and Aging*, 17, 662-676.

Vaden, K.I., Hickok, G.S., & Halpin, H.R. (2009). *Irvine Phonotactic Online Dictionary*, Version 1.4 3.

White, K. K., & Abrams, L. (2002). Does priming specific syllables during tip-of-the-tongue states facilitate word retrieval in older adults? *Psych & Aging*, 17, 226-235.