Published in final edited form as:

Brain Res. 2016 March 15; 1635: 71–85. doi:10.1016/j.brainres.2016.01.014.

Phonological code retrieval during picture naming: Influence of consonant class

Alycia Cummings^a, Amebu Seddoh^a, and Brianna Jallo^a

^aUniversity of North Dakota, Department of Communication Sciences and Disorders

Abstract

Investigations of the time course of various stages of lexical processing have indicated either early or late onset of brain activation for phonological code retrieval. The basis of the differential findings is unclear, but factors related to segmental phonology appear to be part of it. The purpose of the present study was to determine whether phonological encoding is influenced by consonant type. Undergraduate students were presented pictures of common and familiar objects to name. Each picture label had an initial liquid (/l/, /ɹ/) or a stop (/b/, /d/) consonant. Accuracy of picture naming was high and comparable for the two stimulus sets. However, words beginning with liquids elicited larger N2 ERP responses than did those with initial stops. Cluster permutation analysis indicated that the ERP responses elicited by words in the two stimulus sets differed between 293 ms and 371 ms post picture onset. These findings point to a late onset of phonological code retrieval. They have implications for segmental phonology and/or motor planning and execution of speech.

Keywords

event-related potentials (ERP); lexical processing; picture naming; segmental phonology; distinctive features; adult psycholinguistics

1. INTRODUCTION

Lexical processing is one of the most studied phenomena in adult psycholinguistics. Participants in these studies are often presented pictures to name. While picture naming most likely involves both serial and interactive feedback processing, there is nearly universal agreement that it is a staged process in which lexical concepts are generally identified before corresponding phonemes and phonological forms are retrieved (Caramazza & Miozzo, 1997; Levelt, Roelofs, & Meyer, 1999; Levelt, Praamstra, Meyer, Helenius, & Salmelin, 1998; Sahin, Pinker, Cash, Schomer, & Halgren, 2009; Schmitt, Münte, & Kutas, 2000).

Contact Author: Alycia Cummings, 290 Centennial Drive, Stop 8040, Grand Forks, ND 58202-8040, alycia.cummings@und.edu, Phone: 701.777.3725.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of Interest Statement:

The authors have no conflict of interest to disclose.

1.1 Lexical processing stages and their timing

Levelt, Roelofs and Meyer (1999) identified four sequential stages of lexical processing: (1) Conceptual preparation (lexical activation and selection for expression); (2) Lemma retrieval (identification, classification and selection of lexical target); (3) Form encoding, which involves three separate sub-stages including (3a) phonological code retrieval (accessing of individual segments or phonemes that constitute the lexical item), (3b) syllabification (combining the individual segments into syllables), (3c) phonetic encoding (motor programming of individual syllables); and (4) Articulation or production of the word. Although the Levelt-Roelofs-Meyer model (henceforth LRM model) is one of the most influential accounts of lexical processing, timelines for its various components are unspecified. Moreover, investigations of this aspect of the model resulted in only mixed findings.

Early studies devoted to this line of research were inferential analyses of existing behavioral, neuroimaging and electrophysiological data. For example, Indefrey and Levelt (2004) completed a meta-analysis of 82 word-production and 26 auditory-perceptual studies that directly and/or indirectly targeted one or more stages of the LRM model. Their observations suggest that conceptual processes might take place from 0 to about 150–175 ms following picture presentation. Lexical access, phonological code retrieval, and syllabification, on the other hand, were noted to have occurred between 175 and 250 ms, 250 and 333 ms, and 330 ms and 455 ms, respectively. Indefrey (2011) similarly estimated that lexical access might occur around 200 ms, whereas phonological code retrieval, syllabification, and the upper boundary of phonetic encoding take place around 275 ms, 355 ms and 455 ms post-picture onset, respectively.

These data point to late onset of phonological code retrieval. However, it is important to note that phonetic encoding might start as soon as the first phonological syllable has been processed. This possibility implies that there might be an overlap in the time course for phonological and phonetic processes especially in multi-syllabic words. In their recent magnetoencephalography (MEG) study, Miozzo, Pulvermüller, and Hauk (2014) similarly found that phonological and semantic processing occurred simultaneously in the left hemisphere around 150 ms (but see also Valente & Laganaro, 2015).

This finding indicates that phonological code retrieval may actually occur earlier than previous studies suggest (Indefrey, 2011; Indefrey & Levelt, 2004), but it is inconsistent with the findings of some recent investigations. For example, Perret and Laganaro (2012) examined the ERP time course of writing and speaking in 22 adult French speakers. The participants were presented 120 pictures in each condition to name. During the speaking condition, they were told to name the picture verbally as rapidly and as accurately as possible. Similarly in the handwritten condition, they had to write the name of the picture quickly and accurately. No differences were observed in outcomes for the two conditions up to 260 ms post-picture onset. However, different electrophysiological spatial configurations in the two naming conditions were observed between 260 and 600 ms. Specifically, speaking which was the only task that required encoding of phonemes elicited larger N2-type responses than writing, suggesting that it may have directly accounted for the divergence in the outcomes for the two conditions.

A similar finding was reported by Laganaro, Python, and Toepel (2013). These investigators examined ERP responses during picture naming and word reading in 8 adults with impaired phonological-phonetic encoding abilities and 16 healthy controls. ERP amplitudes and global topographies began to differ for the two subject groups around 330 and 400 ms during the reading and the picture-naming tasks, respectively. Since the disordered participants were phonologically impaired, it is reasonable to assume that they would experience difficulty in performing any phonology related task. The divergence of their outcomes from normal between 330 and 400 ms may therefore be an indication that this interval was the period during which phonological-phonetic encoding occurred in the healthy participants.

It appears that the differential findings indicative of early versus late onset of phonological code retrieval might have part of their basis in methodological factors. Miozzo et al. (2014), whose data indicated early onset effect, attributed their results to sensitivity of the technique employed - application of multiple linear regression analysis to MEG data. Previous ERP studies typically used phoneme monitoring Go/No Go tasks, which elicit N2/N200 or lateralized readiness potential (LRP). Go/No Go tasks provide important initial information about accessibility of phonemes, but they also involve interplay of other cognitive processes that could potentially confound results. For example, these tasks require participants to judge whether a specific sound begins a word or a syllable. Their successful performance therefore depends in part on ability to process multiple information such as those on which phonological and semantic decisions are based (Camen, Morand, & Laganaro, 2010; Hanulová, Davidson, & Indefrey, 2011; Jescheniak, Schriefers, Garrett, & Friederici, 2002; Neumann, Obler, Gomes, & Shafer, 2009; Rodriguez-Fornells, Schmitt, Kutas, & Münte, 2002; Schmitt et al., 2000; Turenout, Hagoort, & Brown, 1997, 1998; Zhang & Damian, 2009). Thus, their reliability as a source of information about the time course of phonological code retrieval is unclear.

Studies show that consonants have phonological properties (Luche et al., 2014) that make them enhance lexical processing more than do vowels (Creel, Aslin, & Tanenhaus, 2006; Hochmann, Benavides-Varela, Nespor, & Mehler, 2011; Nespor, Peña, & Mehler, 2003). For example, Havy, Serres, and Nazzi (2014) found that pseudo-words that differed from distractors by one consonant were easier for their French speaking adult subjects to identify compared to similar stimuli that differed by one vowel. Consonant bias has also been observed in ERP N400 responses elicited by lexical decision task (Carreiras, Vergara, & Perea, 2007). Furthermore, it has been demonstrated that primes, which share the same consonant with a target facilitate word recognition better than primes made up of the same vowel as the target (e.g., Lee, Rayner, & Pollatsek, 2002; New & Nazzi, 2014).

Similar to these data, observed behavioral and/or electrophysiological responses are often dependent on type of stimulus and its presentation to participants (Abdel Rahman & Sommer, 2003; Miozzo et al., 2014). Examination of effects of different lexical segmental compositions on behavioral and electrophysiological data might therefore yield information that can provide clues about the time course of phonological code retrieval. We explored this hypothesis by using an approach grounded in distinctive feature theory (Jakobson, Fant, &

Halle, 1952; Ladefoged, 2007). Below is a brief outline of the theoretical background to this approach.

1.3 The present study

It is believed, at least in generative phonology, that distinctive features form the functional basis of phonological systems (e.g., Chomsky & Halle, 1968). They permit categorization of speech sounds that undergo the same phonological rule. The sounds that belong to the same phonological category are considered to form a natural class. Major class features include manner and place of articulation. Manner features (used mostly for consonants) delineate the interaction of articulators for the production of a sound. For example, they distinguish liquids from oral stops. Liquids are vowel-like sounds that are produced with a narrower vocal tract constriction than true vowels (Bleile, 2014; Shriberg & Kent, 2013). By contrast, the production of oral stops is characterized by a complete closure and consequent temporary interruption of airflow at a point in the vocal tract. The interruption results in built-up pressure followed by release of the closure. There are two liquids (lateral /l/ and rhotic /r/) and 6 oral stops (/p, b, t, d, k, g/) in English.

Related to phonological naturalness is the notion of markedness. A consonant is considered marked if it is complex or has irregular distributional pattern in a language. Marked sounds are presumed to require the storage of more distinctive features in their phonological representations compared to simple or regular sounds, which are considered to be unmarked. Liquids are complex or marked sounds. They are acquired late by children presumably because they are difficult to produce (Bleile, 2014; Goldman & Fristoe, 2000; Maddieson, 1984; Poole, 1934; Sander, 1972; Shriberg, 1993; Shriberg & Kent, 2013; Templin, 1957; Wellman, 1931). They also have limited distribution in the world's languages (Gierut, 2001, 2007; Greenberg, 1975). By contrast, oral stops are unmarked in part because nearly every language has at least one of them. They are relatively easy to produce and are typically acquired early by children.

The comparative features for liquids and oral stops suggest that these speech sounds might affect lexical processing differentially. Words with initial liquids may be more difficult to access and produce during picture naming compared to those beginning with oral stops. The purpose of the present study was to determine whether consonant class influences phonological code retrieval. If this stage of lexical processing is affected by consonant type, then participants should demonstrate electrophysiological and/or behavioral differences for words beginning with liquids and stops. Specifically, liquids should elicit larger ERP responses than stops since they are phonologically complex or marked. Also, given that the complexity of a consonant has implication for motor planning and execution, words beginning with liquids should require longer response times compared to those with initial stops.

2. RESULTS

2.1 Behavioral Results

Table 1 shows accuracy and verbal response time data for the subject and the item analyses.

2.1.1 Object Naming Accuracy—Participants labeled pictures beginning with liquids and stops with a high level of accuracy. There was no significant difference between the accuracy of labels beginning with the two consonant types in the subject (F_1) analysis ($p > .20$) or the item (F_2) analysis ($p > .66$).

2.1.2 Verbal Response Time—The subject analysis indicates that object labels beginning with stops elicited faster verbal response times compared to those with initial liquids ($F_1(1,33) = 6.930$, $p < .02$, $\eta^2 = .174$). This effect was not observed in the item analysis ($p > .16$).

2.2 ERP Results

Table 2 displays means and standard deviations for amplitude measures (in microvolts) in 9 electrode groups for P2 and N2 peaks. Across the Posterior/Occipital electrodes, a P1-N1-P2 peak sequence commonly associated with visual stimulus presentation was observed. More central and anterior electrodes presented with an N1-P2-N2-P3 peak sequence. Detailed neural activation data for P2 and N2 amplitudes are presented below.

2.2.1 P2 amplitude—There was no significant difference in P2 amplitudes for words beginning with liquids and stops ($p > .09$). On the other hand, an effect of electrode group location was observed ($F(8,264) = 32.964$, $p < .0001$, $\eta^2 = .500$). A general trend of more negative-going ERP responses was observed over frontal electrode locations. By contrast, more positive-going ERP responses occurred in posterior scalp areas. Post-hoc pairwise comparisons revealed no difference between outputs for the left frontal, fronto-central, and right frontal electrode groups. These electrode groups indicated more negative responses compared to the left central and the right central electrodes, which were subsequently more negative than the centro-parietal group. The centro-parietal electrode group also elicited more negative ERP responses than did the left parietal and right parietal groups. The occipital electrodes elicited the most positive-going ERP responses (i.e., $LF = FC = RF < LC = RC < CP < LP = RP < O$). No Consonant class $\times$ Electrode group interaction was found.

ERP waveforms for all 42 electrodes from which amplitude measures were made and midline electrodes for P2-N2-P3 peaks are displayed in Figures 1 and 2, respectively.

2.2.2 N2 Amplitude—Picture labels beginning with liquids elicited larger (i.e., more negative) N2 responses than did those with initial stops ($F(1,33) = 4.965$, $p < .04$, $\eta^2 = .131$). An effect of electrode group location was also observed ($F(8,264) = 53.983$, $p < .0001$, $\eta^2 = .621$). Electrode group differences observed for N2 responses were similar to those demonstrated for P2 (i.e., $LF = FC = RF < LC = RC < CP < LP = RP < O$). Furthermore, a general trend of more negative-going ERP responses was observed over frontal electrode locations whereas more positive-going ERP responses occurred over posterior scalp areas. No Consonant class $\times$ Electrode group interaction was found.

2.2.3 Cluster Permutation Analyses

2.2.3.1 Consonant Class: Cluster-level mass permutation procedure encompassing both early and late onset phonological processing time windows (150–400 ms) was applied to the data in order to determine the precise time course of phonological code retrieval of picture names beginning with liquids and stops. The results of this test are displayed in a raster diagram in Figure 3. There was one significant broadly distributed cluster that began at 293 ms and ended at 371 ms. Thus, all differences were found to be within the late onset phonological processing time window. No difference during the early onset time period was observed. Consonant class differences were fairly prominent across the entire scalp (i.e., they were not localized to one hemisphere or midline). This was consistent with the amplitude results that indicated no Consonant class x Electrode group interaction.

2.2.3.2 Word Frequency: Cluster permutation analyses of the frequency data showed varied results. The high frequency and the combined high versus low frequency analyses revealed no significant clusters. However, visual inspection of the high frequency data indicates that the liquids elicited slightly larger N2 responses than did the stops (Figure 4A). By contrast, the low frequency data showed two significant broadly distributed clusters. One began at 187 ms and ended at 254 ms; the second began at 280 ms and ended at 400 ms, which is the boundary of the analysis window (Figure 5). Liquids in the low frequency data elicited larger N2 responses than did stops in both time windows (Figure 4B). Thus, consonant class differences were found across nearly the entire 250 ms analysis window, which represents both the proposed early (Miozzo et al., 2014) and late (Indefrey, 2011; Indefrey & Levelt, 2004) phonological processing time windows.

3. DISCUSSION

This study examined electrophysiological and behavioral indices of phonological code retrieval during picture naming in English-speaking adults. The initial consonant of each picture label was either a liquid (lateral /l/ or rhotic /r/) or a stop (/b/ or /d/). Unlike accuracy (and to some extent reaction time) measures, ERP N2 responses differed significantly according to stimulus initial consonant class. Picture labels with initial liquids elicited larger N2 responses compared to those with initial stops for the time period of 293–371 ms post-picture onset (no differences were found prior to and after this time period). These electrophysiological outcomes suggest that consonant class may influence the retrieval of phonological codes.

3.1 Neural evidence for sound complexity

Large ERP responses may be a reflection of a high level of neural involvement during task performance (Otten & Rugg, 2005). The finding that N2 responses for picture labels beginning with liquids were larger than those for words with initial stops provides neural evidence in support of the notion that phonemes vary in terms of their distinctive feature composition and complexity. The retrieval of words beginning with liquids may require recruitment of more neural resources compared to the case for words with initial stops.

This interpretation is supported by the results of the frequency analysis. Low frequency words with initial liquids elicited larger N2 responses than did their stop counterparts over two extended time periods (187–254 ms and 280–400 ms). Results of the analysis of the combined data for labels beginning with the two consonant types, on the other hand, showed no difference between high versus low frequency words. These findings are consistent with the view that phonological encoding of liquids requires more neural resources compared to the case for stops. They suggest that consonant class rather than word frequency may have driven the observed response differences. However, given that phonetic processing overlaps with phonological encoding, the current data may also be related to motor planning and/or execution of speech.

It is possible that monosyllabic and multisyllabic words elicited different ERP effects. This possibility was examined by comparing the one-syllable words to the multisyllabic words (two- and three-syllable) in both stimulus types. Repeated measure, two-tailed cluster-based permutation tests revealed no significant syllable length difference for the words in each stimulus category, although there were some variations in the ERP responses due to the small number of trials (Figure 6). When the one-syllable words from each stimulus category were combined and compared with all of the multisyllabic words, the ERP waveforms were nearly identical, suggesting that syllable length differences may have no influence on the phonological processing time windows. The absence of syllable length effect supports the decision to collapse ERP responses across all words. It suggests that the phonological encoding differences in the overall ERP responses might be an effect of word initial consonant rather than syllable size.

The present ERP data therefore provide the first documented evidence in support of the conceptualization of the phoneme as a bundle of distinctive features. These data suggest that phonological representation might consist of stored distinctive features, and that complex or marked phonemes might require more mental effort for their production and/or perception compared to the case for their simple or unmarked counterparts.

3.2 Early vs. late phonological code retrieval

Two different ERP measures provide converging evidence for the timing of phonological processing. In the peak amplitude measures, significant differences were observed in the late N2 peak, but not the earlier P2 peak. Moreover, the cluster permutation analysis indicated an approximately 80 ms time period during which words beginning with liquids and stops elicited different ERP responses (293–371 ms post-picture onset). Thus, the observed timing differences are consistent with the late onset phonological processing effect proposed by Indefrey and Levelt (2004, 2011).

An alternative interpretation is suggested by the long verbal response times. These data may be an indication of extended time course for each stage of the processing of the stimuli including phonological code retrieval. The long time course can result from prolonged overall neural activation. This possibility implies that the processing of phonological information within the time window of 293–371 ms is relatively early rather than late.¹

¹We owe this interpretation to an anonymous reviewer.

One problem with this interpretation is the implication that ERP outcomes are dependent on the timing of verbal responses. Phonological form encoding has been shown to be a fairly automatic process (Ferreira & Pashler, 2002). In particular, there is lack of modulation between its time window and the speed of production (Laganaro et al., 2013). The verbal response times demonstrated by participants in some recent picture naming studies (Laganaro, Valente, & Perret, 2012; Perret & Laganaro, 2012; Strijkers, Costa, & Thierry, 2010) are much faster than those in the current study. Yet ERP waveforms displayed by participants in these studies and the current investigation have similar morphology.

It may be that the extended time window for the verbal responses had methodological rather than theoretical basis. Perhaps calibration of the voice key was faulty. A poorly calibrated voice key can lead to large amounts of variation in response times (Plant & Quinlan, 2013). A delay in the transmission of participants' responses could also stem from the computer sound card to which the voice key was connected. Other potential sources of this outcome include stimulus presentation software, the ActiveTwo ERP recording software program, and/or ERPLAB (Luck & Lopez-Calderon, 2012) data analysis software program. Exploration of these possibilities in future research is warranted.

To ascertain whether there was a transmission delay, latencies of participants' verbal responses were measured. A new group of 10 female undergraduate native speakers of English (aged 20–25 years) was recruited to participate in a separate experiment for this purpose. This experiment was necessitated by a need to conduct the measures offline rather than do so online with the voice key. Apart from the calibration problem mentioned above, the voice key has a limited ability to capture verbal response times reliably. It can, for example, be activated by participant activities such as breathing and touching that are not necessarily related to target responses (see also section 3.3 below). Given that signals for different activities are indistinguishable, there is a potential for error when the timing of verbal responses is measured online with voice key.

Similar to the original experiment, the new participants were presented the stimulus pictures to label. However, unlike the case for the previous approach, each picture onset was marked with an audible click, which was audio-recorded together with the participants' verbal responses. The temporal gap between the onset of the click and the response was then hand-measured. Results showed that on average, median response times for stop-initial and liquid-initial words were 745 (SD = 171) ms and 740 (SD = 191) ms, respectively, which were 381 ms and 416 ms faster than those calculated by ERPLAB (Luck & Lopez-Calderon, 2012).

These findings not only support the possibility of a delay in response transmission in the original experiment, but they are also consistent with the latencies reported by previous studies (Laganaro et al., 2012; Perret & Laganaro, 2012; Strijkers et al., 2010). It is important to note, however, that unlike the case for ERP data that are generated automatically, verbal response latency in picture naming may be influenced by a number of factors including attention and motor control. Whether the relatively large variability in the current data is related to task difficulty or the mechanism(s) underlying participant performance is unclear. Uncertainties like these call for caution in interpreting verbal

response time data such as the above. The need for exploration of this aspect of picture naming in future research cannot be overemphasized.

3.3 Behavioral effects of sound complexity

It was hypothesized that words beginning with liquids would require longer response times than those with initial stops. The subject based analysis indicated that verbal response times for picture labels beginning with liquids were significantly slower than those for words with initial stops. One possibility is that the complexity of liquids made it difficult for participants to quickly retrieve or produce words beginning with these consonants. Although this interpretation is consistent with the ERP data, its validity is questionable given the outcome of the item analysis. Unlike the subject-based analysis, the item analysis indicated no difference between verbal response times for words beginning with stops and liquids. This discrepancy suggests that rather than being a reflection of varied levels of effort in encoding stops versus liquids, the differential verbal response times for the words beginning with the two consonant types might be an indication of individual differences that are specific to the stimuli used.

This possibility is supported by outcomes of the voice key measures. Although the Cedrus SV-1 voice key has capacity to record responses with 1 ms resolution, its overall timing accuracy can be affected by participant differences. In the current study, participants who were soft-spoken exhibited fewer than expected voice key responses as some of their productions failed to register. On the other hand, those whose speech productions were staccato-like showed many voice key responses that may have been triggered by each consonant onset and/or release.

3.4 Clinical implication

The current findings have implications for clinical speech-language pathology, particularly speech sound production impairment. If it is confirmed that phonological retrieval of different sounds yields differential neural response patterns, that might provide a basis for understanding the scientific and clinical nature of this disorder. Behavioral and/or electrophysiological tests could be administered pre- and post-treatment to explore neural changes that occur secondary to therapy. Such evaluations could possibly lead to identification of different sub-groups of patients with the disorder. For example, some patients may have difficulty retrieving phonological information and/or its production. Others may have only speech production problems. Studies of adults and children with different kinds of communication disorders are warranted to explore these possibilities.

3.5 Limitations

To our knowledge, this is the first study to examine the effect of consonant complexity on phonological code retrieval. The experimental task was designed to be simple enough to allow its successful completion using minimal effort. Consequently, it has a limited number of stimuli compared to previous MEG and ERP studies (Aristei, Melinger, & Abdel Rahman, 2011; Costa, Strijkers, Martin, & Thierry, 2009; Laganaro, 2014; Laganaro & Perret, 2011; Levelt et al., 1998; Maess, Friederici, Damian, Meyer, & Levelt, 2002; Perret & Laganaro, 2012; Strijkers et al., 2010). The relatively small number of stimulus words (40

with initial liquids and 42 with initial stops) could have led to a poor signal-to-noise ratio. On the other hand, participant fatigue could be a factor if the number of stimuli was increased. Even with only 116 words (82 stimuli and 34 filler trials), the participants required up to fifteen minutes to complete the task. The recruitment of 34 participants was intended to address the potential signal-to-noise problem (there were 10–24 more participants in this study compared to those listed above). However, whether this approach resolved the problem completely is unclear.

Also, in line with the objective of determining whether consonant class influences phonological retrieval, object labels were controlled across multiple lexical and phonological variables, including age of acquisition, density, and syllable length. On the other hand, phonemes in the picture labels except those at word initial positions were uncontrolled. It is therefore possible that our results were affected by phonological properties of the consonants in non-initial positions. The only way to explicitly control for effects of all phonemes in a naming task would be to use minimal pairs such as *rake* vs. *lake* or *lair* vs. *bear*. However, this amount of control could induce lexical and or syntactic effects due to factors such as differences in word frequency, neighborhood density, and/or age of acquisition, not to mention the possible difficulty with accurate representation of some of these concepts in a single, simple picture.

4. EXPERIMENTAL PROCEDURE

4.1 Participants

Thirty-four female native speakers of (American) English aged 19–23 years participated in the study. All of them were undergraduate students at the University of North Dakota, U.S.A. They had normal or corrected-to-normal vision, and none had a history of speech, language, and/or hearing impairment. Each signed informed consent in accordance with the University of North Dakota Human Research Protection Program.

4.2 Stimuli

The stimuli were 116 pictures of colored cartoon drawings of objects. All pictures were 345 × 270 pixels. They were selected based on the result of a prior survey that sought to identify labels most used for various objects. The survey involved presentation of 150 pictures individually on an overhead projector to a group of 30 native speakers of (American) English, who were also undergraduate students. The students were asked to write down the label they considered most representative of the name for each pictured object. The most dominant response for each picture was selected. Ambiguous pictures for which two or more participants provided an alternative name were excluded from consideration. This approach led to the selection of 82 stimulus pictures and their labels (see appendix). There were also 34 filler trials that were not included in the analyses. The consonant initiating the target label for each picture was either a liquid (/l/, /ɹ/) or a stop (/b/, /d/). Forty of the target picture labels had initial liquids (20 with /l/ and 20 with /ɹ/) and 42 had initial stops (27 with /b/ and 15 with /d/).

Various phonological and lexical characteristics of the target labels were controlled as closely as possible. Words with initial liquids had comparable number of syllables as those beginning with stops ($p > .07$). Age of acquisition (AoA), the period when a word enters a child's vocabulary (Sosa & Stoel-Gammon, 2012), was examined based on a database of over 30,000 English words (<http://crr.ugent.be/archives/806>) created by Kuperman and colleagues (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012). There was no significant difference ($p > .13$) in AoA for the words beginning with liquids and stops. Neighborhood density or the number of phonological neighbors was also examined. Phonological neighbors refer to words that differ from one another in terms of addition, deletion, or substitution of a single phoneme (Luce & Pisoni, 1998; Sosa & Stoel-Gammon, 2012). Stimulus words beginning with liquids had comparable number of phonological neighbors as those with initial stops ($p > .29$).

In contrast to the above, words beginning with stops had higher log10 word frequencies (log10WF) than those with initial liquids ($t(80) = -2.684$, $p < .01$). Values for the log10WF measure were gathered from the SUBTLEX_{US} corpus (Brysbaert & New, 2009). Briefly, the frequency count (FREQcount) of the SUBTLEX_{US} corpus refers to the number of times a word appears in the corpus of 51 million spoken words. The log10WF value was based on $\log_{10}(\text{FREQcount}+1)$. Table 3 displays the characteristics of the stimuli used.

4.3 Procedure

Prior to the actual experiment, participants were trained to name the pictures verbally without preceding their productions with other words or sounds such as “a”, hmmm.... They were taught to say, “don’t know” when they failed to remember a word or recognize an object. Six pictured objects that were not part of the actual experimental stimuli were used for the training.

Participants were tested in a well-lit, quiet sound-controlled booth. They wore a voice-activated microphone (Cedrus SV-1; <http://cedrus.com/sv1/>) around their necks. The microphone was placed approximately three inches in front of each participant's mouth, and it recorded the timing of the onset of each verbal response. The picture stimuli were presented to participants using Presentation® software package (www.neurobs.com). Measures of verbal response times were also conducted using this software.

Participants were shown the stimuli, one at a time, on a computer screen in front of them. They were asked to label each picture as it appeared as quickly and accurately as possible. All pictures were presented in the same order to all participants. They remained on the screen until the participant provided verbal labels for them, or for up to 4000 ms, whichever came first. A blank screen containing a fixation crosshatch (“+”) was presented between trials with an inter-stimulus interval of 1500 ms (Figure 7). The experiment lasted for approximately 15 minutes per participant. Each participant's session was audio-recorded for off-line phonetic transcription and accuracy analysis. This was in order to ensure that all on-line button presses accurately represented the participant's verbal responses (see below).

4.4 Behavioral Data Analysis

Response accuracy was recorded on-line immediately following each label for each participant via an experimenter-initiated button response (Cedrus RB-530, <http://cedrus.com/responsepads/rb530.htm>). Following prior object naming procedures (Szekely et al., 2005), each verbal label was assigned to one of the following accuracy categories: (1) Correct responses (exact target object label); (2) Alternate responses (synonyms, hypernyms/superordinate words, hyponyms); (3) Incorrect responses; (4) “Don’t know” responses; (5) Mispronounced words (e.g., saying “wope” for “rope”).

Reaction time was calculated from the onset of the visual stimulus to the participant’s voice onset. Acceptable reaction times were those that fell below 4000 ms. This temporal window is consistent with those for previous studies (Aristei et al., 2011; Costa et al., 2009; Laganaro, 2014; Laganaro & Perret, 2011; Levelt et al., 1998; Maess et al., 2002; Strijkers et al., 2010). Pictures eliciting responses longer than 4000 ms were not considered in data analysis. Since there was a wide variation in reaction times, median reaction times were used. Naming accuracy and verbal response times were calculated for correct trials only. The effect of word-initial consonant class on naming accuracy and verbal response time was assessed in two separate ANOVAs with subject (F_1) and item (F_2) as random factors. Both analyses were conducted on the correctly named trials only.

4.5 EEG Recording and Averaging

Sixty-six channels of continuous EEG (DC–128 Hz) were recorded using ActiveTwo data acquisition system (Biosemi, Inc, Amsterdam, Netherlands) at a sampling rate of 256 Hz. This system provides “active” EEG amplification at the scalp that substantially minimizes movement artifacts. The amplifier gain on this system is fixed, allowing ample input range (–264 to 264 mV) on a wide dynamic range (110 dB) Delta-Sigma (Σ) 24-bit AD converter. Sixty-four channel scalp data were recorded using electrodes mounted in a stretchy cap according to the International 10–20 system. Two additional electrodes were placed on the right and left mastoids. Eye movements were monitored using FP1/FP2 (blinks) and F7/F8 channels (lateral movements, saccades). During data acquisition, all channels were referred to the system’s internal loop (CMS/DRL sensors located in the centro-parietal region), which drives the average potential of a subject (the Common Mode voltage) as close as possible to the analog-digital converter reference voltage (the amplifier “zero”). The DC offsets were kept below 25 microvolts at all channels. Off-line, data were re-referenced to the average of left and right mastoid tracings.

Prior to data averaging, sporadic artifact rejection of the continuous EEG was completed using EEGLAB (Delorme & Makeig, 2004). This involved marking and rejecting the time periods during which sporadic artifacts occurred. Sources of the artifacts included random head movements, muscle movements related to speaking, excessive electrode activation stemming from pressing the head into the back of the chair, etc. After sporadic artifact rejection, independent-component analysis (ICA) (Jung et al., 2000) was completed so that the experimenters could identify eye blink and saccade components. These components were then deleted from the continuous EEG. The remaining artifactual trials due to amplifier blocking as well as muscle and overall body movements were rejected from further analyses

using the simple voltage threshold measure in ERPLAB (Luck & Lopez-Calderon, 2012). Voltage limits were set at -100 to 100 microvolts. Participants were excluded from the study if more than 25% of their data exceeded this threshold.

Epochs containing 100 ms pre-auditory stimulus and 900 ms stimulus time were baseline-corrected with respect to the pre-stimulus interval and averaged by stimulus type (i.e., words beginning with stops versus liquids). The data were low-pass filtered at 30 Hz and high-pass filtered at 0.5 Hz using two-way least squares FIR filters. On average, the remaining individual data contained 36 ($SD = 3$) word-initial stop and 32 ($SD = 5$) liquid trials.

4.6 ERP Measurements

4.6.1 Mean Amplitude Measurements—Only correct trials were included in the ERP mean amplitude analyses. Since the timing of phonological processing (or specifically phonological code retrieval) was of interest in this experiment, mean amplitudes were measured across two time windows broadly encompassing two ERP peaks: 150–275 ms (i.e., the P2 peak) and 275–400 ms (i.e., the N2 peak). The timing of the mean amplitude measurements was based on the proposed early (Miozzo et al., 2014) and late (Indefrey, 2011; Indefrey & Levelt, 2004) onset phonological processing timelines discussed earlier. The timing of these two peaks (150–400 ms) allowed for an evaluation of both of these proposals.

The electrodes for both peak analyses were chosen based on groupings in a previous picture naming study (Strijkers et al., 2010). Forty-two electrodes were put into 9 different groups based on scalp location (Figure 8) as follows. (1) Left Frontal (LF): F5, F3, FC5, FC3; (2) Fronto-Central (FC): F1, Fz, F2, FC1, FCz, FC2, C1, Cz, C2; (3) Right Frontal (RF): F6, F4, FC6, FC4; (4) Left Central (LC): C5, C3, CP5, CP3; (5) Centro-Parietal (CP): CP1, CPz, CP2, P1, Pz, P2; (6) Right Central (RC): C6, C4, CP6, CP4; (7) Left Parietal (LP): P5, P3, PO7, PO3; (8) Occipital (O): O1, Oz, O2; (9) Right Parietal (RP): P6, P4, PO8, PO4.

Mean amplitudes from each of the 42 electrodes were calculated separately for the P2 and N2 responses. The 42 electrodes were then separated into their respective location-specific groups. A median amplitude measurement for each of the nine electrode groups was then calculated. This calculation was done using the mean amplitudes of the electrodes in each group. Repeated measure ANOVAs were completed separately for both time windows (P2, N2) using the median group amplitudes: Consonant class (liquid, stop) $\times$ Electrode group (1–9). Bonferroni corrections were applied when appropriate. Partial eta squared (η^2) effect sizes were noted for all significant effects and interactions.

4.6.2 Cluster Mass Permutation Tests

4.6.2.1 Consonant Class Analysis: A repeated measure two-tailed cluster-based permutation test was conducted on the ERPs from the two consonant classes (word initial liquids and stops) using a family-wise alpha level of 0.05 (Bullmore et al., 1999; Groppe, Urbach, & Kutas, 2011a). This was in order to determine whether there were reliable differences between the ERPs for word initial liquids and stops. All time points between 150 and 400 ms at 43 scalp electrodes (all 42 electrodes described above in the mean amplitude

analysis, as well as POz) were included in the test (i.e., 2795 total comparisons). T-tests were performed for each comparison using the original data and 2500 random within-participant permutations of the data. For each permutation, all t-scores corresponding to uncorrected p-values of 0.05 or less were formed into clusters.

Electrodes within about 5.44 cm of one another were considered spatial neighbors, and adjacent time points were considered temporal neighbors. The sum of the t-scores in each cluster was the "mass" of that cluster. The most extreme cluster mass in each of the 2501 sets of tests was recorded and used to estimate the distribution of the null hypothesis (i.e., no difference between conditions). The permutation cluster mass percentile ranking of each cluster from the observed data was used to derive p-values assigned to each member of the cluster. T-scores that were not included in a cluster were given a p-value of 1.

This permutation test analysis provides a better spatial and temporal resolution than conventional ANOVAs while maintaining weak control of the family-wise alpha level (i.e., it corrects for the large number of comparisons). Moreover, the cluster mass statistic was chosen for this permutation test because it has been shown to have relatively good power for broadly distributed ERP effects (Groppe, Urbach, & Kutas, 2011b; Maris & Oostenveld, 2007). To estimate the distribution of the null hypothesis, 2500 permutations were used, which was more than twice the number recommended for a family-wise alpha level of 0.05 (Manly, 2006). In other words, the cluster mass permutation analysis allowed for a full examination of the two proposed timeline windows for phonological processing. This analysis enabled identification of differences between the words with initial liquids and stops at any point in time during the 150–400 ms (i.e., 250 ms) analysis window. Thus, the high temporal resolution of this analysis could be used to identify a specific time period during which phonological processing occurred.

4.6.2.2 Word Frequency Analysis: Since the stimulus labels with initial stops had higher log10WF values compared to those with initial liquids, there was a potential for ERP results to be influenced by word frequency. To explore this possibility, a consonant class x word frequency analysis was completed. In view of the small number of trials per participant, the stop-initial and the liquid-initial words were categorized into low and high frequency groups based on their log10WF values (see Appendix). The categorization was based on the median log10WF value of each consonant class. This ensured that the elements of both categories had equal numbers of trials. Low and high frequency labels with initial stops had log10WF values ranging from 1.65 to 2.89 and 2.99 to 3.99, respectively. The corresponding values for labels with initial liquids were 1.08 to 2.65 and 2.67 to 3.67. Cluster mass permutation tests were conducted to compare each frequency set for words with initial stops with the corresponding set for words with initial liquids. Additionally, the high frequency words with the two consonant types were combined. A similar combination involving the low frequency words was made and compared with the combined high frequency words. Similar to the cluster mass permutation test for the consonant class data (section 4.6.2.1), all time points between 150 and 400 ms at 43 scalp electrodes were included in each test (i.e., 2795 total comparisons).

Acknowledgments

This research was supported by NIH grant numbers R15DC013359 (from the National Institute On Deafness And Other Communication Disorders) and C06RR022088 (from the National Center for Research Resources) awarded to the first author. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We would like to thank two anonymous reviewers for their valuable comments on previous drafts of the paper. We also appreciate the help of Janet Babchishin, Megan Nauman, Meghan Macaulay, Courtney Rowan, Sheila Cassidy, Kristyne Hass, Anna Morris, and Michaela Johnson in testing participants and data preparation. The authors do not have any financial or non-financial relationships relevant to the content of this manuscript.

REFERENCES

- Abdel Rahman R, Sommer W. Does phonological encoding in speech production always follow the retrieval of semantic knowledge?: Electrophysiological evidence for parallel processing. *Cognitive Brain Research*. 2003; 16(3):372–382. [http://doi.org/10.1016/S0926-6410\(02\)00305-1](http://doi.org/10.1016/S0926-6410(02)00305-1). [PubMed: 12706217]
- Aristei S, Melinger A, Abdel Rahman R. Electrophysiological chronometry of semantic context effects in language production. *Journal of Cognitive Neuroscience*. 2011; 23(7):1567–1586. <http://doi.org/10.1162/jocn.2010.21474>. [PubMed: 20515409]
- Bleile, KM. The late eight. 2nd. San Diego, CA: Plural Pub; 2014.
- Brysbaert M, New B. 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*. 2009; 41(4):977–990. <http://doi.org/10.3758/BRM.41.4.977>. [PubMed: 19897807]
- Bullmore ET, Suckling J, Overmeyer S, Rabe-Hesketh S, Taylor E, Brammer MJ. Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. *IEEE Transactions on Medical Imaging*. 1999; 18(1):32–42. [PubMed: 10193695]
- Camen C, Morand S, Laganaro M. Re-evaluating the time course of gender and phonological encoding during silent monitoring tasks estimated by ERP: serial or parallel processing? *Journal of Psycholinguistic Research*. 2010; 39(1):35–49. <http://doi.org/10.1007/s10936-009-9124-4>. [PubMed: 19644758]
- Caramazza A, Miozzo M. The relation between syntactic and phonological knowledge in lexical access: evidence from the “tip-of-the-tongue” phenomenon. *Cognition*. 1997; 64(3):309–343. [PubMed: 9426505]
- Carreiras M, Vergara M, Perea M. ERP correlates of transposed-letter similarity effects: are consonants processed differently from vowels? *Neuroscience Letters*. 2007; 419(3):219–224. <http://doi.org/10.1016/j.neulet.2007.04.053>. [PubMed: 17507160]
- Chomsky, N.; Halle, M. The sound pattern of English. New York: Harper & Row; 1968.
- Costa A, Strijkers K, Martin C, Thierry G. The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proceedings of the National Academy of Sciences of the United States of America*. 2009; 106(50):21442–21446. <http://doi.org/10.1073/pnas.0908921106>. [PubMed: 19934043]
- Creel SC, Aslin RN, Tanenhaus MK. Acquiring an artificial lexicon: Segment type and order information in early lexical entries. *Journal of Memory and Language*. 2006; 54(1):1–19. <http://doi.org/10.1016/j.jml.2005.09.003>.
- Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*. 2004; 134(1):9–21. <http://doi.org/10.1016/j.jneumeth.2003.10.009>. [PubMed: 15102499]
- Ferreira VS, Pashler H. Central bottleneck influences on the processing stages of word production. *Journal of Experimental Psychology. Learning, Memory, and Cognition*. 2002; 28(6):1187–1199.
- Gierut JA. Complexity in Phonological Treatment: Clinical Factors. *Language, Speech, and Hearing Services in Schools*. 2001; 32(4):229–241. [http://doi.org/10.1044/0161-1461\(2001/021\)](http://doi.org/10.1044/0161-1461(2001/021)).

- Gierut JA. Phonological complexity and language learnability. *American Journal of Speech-Language Pathology / American Speech-Language-Hearing Association*. 2007; 16(1):6–17. [http://doi.org/10.1044/1058-0360\(2007/003\)](http://doi.org/10.1044/1058-0360(2007/003)). [PubMed: 17329671]
- Goldman, R.; Fristoe, M. Goldman-Fristoe Test of Articulation. 2nd. Minneapolis, MN: Pearson; 2000.
- Greenberg JH. Research on Language Universals. *Annual Review of Anthropology*. 1975; 4(1):75–94. <http://doi.org/10.1146/annurev.an.04.100175.000451>.
- Groppe DM, Urbach TP, Kutas M. Mass univariate analysis of event-related brain potentials/fields I: a critical tutorial review. *Psychophysiology*. 2011a; 48(12):1711–1725. <http://doi.org/10.1111/j.1469-8986.2011.01273.x>. [PubMed: 21895683]
- Groppe DM, Urbach TP, Kutas M. Mass univariate analysis of event-related brain potentials/fields II: Simulation studies. *Psychophysiology*. 2011b; 48(12):1726–1737. <http://doi.org/10.1111/j.1469-8986.2011.01272.x>. [PubMed: 21895684]
- Hanulová J, Davidson DJ, Indefrey P. Where does the delay in L2 picture naming come from? Psycholinguistic and neurocognitive evidence on second language word production. *Language and Cognitive Processes*. 2011; 26(7):902–934. <http://doi.org/10.1080/01690965.2010.509946>.
- Havy M, Serres J, Nazzi T. A consonant/vowel asymmetry in word-for processing: evidence in childhood and in adulthood. *Language and Speech*. 2014; 57(Pt 2):254–281. [PubMed: 25102609]
- Hochmann J-R, Benavides-Varela S, Nespor M, Mehler J. Consonants and vowels: different roles in early language acquisition. *Developmental Science*. 2011; 14(6):1445–1458. <http://doi.org/10.1111/j.1467-7687.2011.01089.x>. [PubMed: 22010902]
- Indefrey P. The spatial and temporal signatures of word production components: a critical update. *Frontiers in Psychology*. 2011; 2:255. <http://doi.org/10.3389/fpsyg.2011.00255>. [PubMed: 22016740]
- Indefrey P, Levelt WJM. The spatial and temporal signatures of word production components. *Cognition*. 2004; 92(1–2):101–144. <http://doi.org/10.1016/j.cognition.2002.06.001>. [PubMed: 15037128]
- Jakobson, R.; Fant, G.; Halle, M. Preliminaries to speech analysis. the distinctive features and their correlates. Cambridge, MA: MIT, Acoustics Laboratory; 1952. Technical Report No. 13
- Jescheniak JD, Schriefers H, Garrett MF, Friederici AD. Exploring the activation of semantic and phonological codes during speech planning with event-related brain potentials. *Journal of Cognitive Neuroscience*. 2002; 14(6):951–964. <http://doi.org/10.1162/089892902760191162>. [PubMed: 12191461]
- Jung TP, Makeig S, Westerfield M, Townsend J, Courchesne E, Sejnowski TJ. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*. 2000; 111(10):1745–1758. [PubMed: 11018488]
- Kuperman V, Stadthagen-Gonzalez H, Brysbaert M. Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*. 2012; 44(4):978–990. <http://doi.org/10.3758/s13428-012-0210-4>. [PubMed: 22581493]
- Ladefoged P. Articulatory features for describing lexical distinctions. *Language*. 2007; 83(1):161–180.
- Laganaro M. ERP topographic analyses from concept to articulation in word production studies. *Frontiers in Psychology*. 2014; 5:493. <http://doi.org/10.3389/fpsyg.2014.00493>. [PubMed: 24904505]
- Laganaro M, Perret C. Comparing electrophysiological correlates of word production in immediate and delayed naming through the analysis of word age of acquisition effects. *Brain Topography*. 2011; 24(1):19–29. <http://doi.org/10.1007/s10548-010-0162-x>. [PubMed: 20938730]
- Laganaro M, Python G, Toepel U. Dynamics of phonological-phonetic encoding in word production: evidence from diverging ERPs between stroke patients and controls. *Brain and Language*. 2013; 126(2):123–132. <http://doi.org/10.1016/j.bandl.2013.03.004>. [PubMed: 23707932]
- Laganaro M, Valente A, Perret C. Time course of word production in fast and slow speakers: a high density ERP topographic study. *NeuroImage*. 2012; 59(4):3881–3888. <http://doi.org/10.1016/j.neuroimage.2011.10.082>. [PubMed: 22079505]

- Lee H-W, Rayner K, Pollatsek A. The processing of consonants and vowels in reading: evidence from the fast priming paradigm. *Psychonomic Bulletin & Review*. 2002; 9(4):766–772. [PubMed: 12613681]
- Levelt Roelofs A, Meyer AS. A theory of lexical access in speech production. *The Behavioral and Brain Sciences*. 1999; 22(1):1–38. discussion 38–75. [PubMed: 11301520]
- Levelt WJM, Praamstra P, Meyer AS, Helenius P, Salmelin R. An MEG study of picture naming. *Journal of Cognitive Neuroscience*. 1998; 10(5):553–567. [PubMed: 9802989]
- Luce PA, Pisoni DB. Recognizing spoken words: the neighborhood activation model. *Ear and Hearing*. 1998; 19(1):1–36. [PubMed: 9504270]
- Luche CD, Poltrock S, Goslin J, New B, Floccia C, Nazzi T. Differential processing of consonants and vowels in the auditory modality: A cross-linguistic study. *Journal of Memory and Language*. 2014; 72:1–15. <http://doi.org/10.1016/j.jml.2013.12.001>.
- Luck, S.; Lopez-Calderon, J. ERPLAB Toolbox (Version 3.0.2.1). Davis: University of California; 2012.
- Maddieson, I. *Patterns of Sounds*. Cambridge University Press; 1984.
- Maess B, Friederici AD, Damian M, Meyer AS, Levelt WJM. Semantic category interference in overt picture naming: sharpening current density localization by PCA. *Journal of Cognitive Neuroscience*. 2002; 14(3):455–462. <http://doi.org/10.1162/089892902317361967>. [PubMed: 11970804]
- Manly, BFJ. *Randomization, Bootstrap and Monte Carlo Methods in Biology*. Third. CRC Press; 2006.
- Maris E, Oostenveld R. Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*. 2007; 164(1):177–190. <http://doi.org/10.1016/j.jneumeth.2007.03.024>. [PubMed: 17517438]
- Miozzo M, Pulvermüller F, Hauk O. Early Parallel Activation of Semantics and Phonology in Picture Naming: Evidence from a Multiple Linear Regression MEG Study. *Cerebral Cortex* (New York, N.Y.: 1991). 2014 <http://doi.org/10.1093/cercor/bhu137>.
- Nespor M, Peña M, Mehler J. On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue E Linguaggio*. 2003; 2:203–229.
- Neumann Y, Obler LK, Gomes H, Shafer V. Phonological vs sensory contributions to age effects in naming: An electrophysiological study. *Aphasiology*. 2009; 23(7–8):1028–1039. <http://doi.org/10.1080/02687030802661630>.
- New B, Nazzi T. The time course of consonant and vowel processing during word recognition. *Language, Cognition and Neuroscience*. 2014; 29(2):147–157. <http://doi.org/10.1080/01690965.2012.735678>.
- Otten, L.; Rugg, M. Interpreting event-related brain potentials. In: Handy, T., editor. *Event-related potentials: A methods handbook*. Cambridge, MA: MIT Press; 2005.
- Perret C, Laganaro M. Comparison of electrophysiological correlates of writing and speaking: a topographic ERP analysis. *Brain Topography*. 2012; 25(1):64–72. <http://doi.org/10.1007/s10548-011-0200-3>. [PubMed: 21863371]
- Plant RR, Quinlan PT. Could millisecond timing errors in commonly used equipment be a cause of replication failure in some neuroscience studies? *Cognitive, Affective, & Behavioral Neuroscience*. 2013; 13(3):598–614. <http://doi.org/10.3758/s13415-013-0166-6>.
- Poole I. Pupil cooperation in speech program activities. *Educ. Meth*. 1934; 13:297–300.
- Rodriguez-Fornells A, Schmitt BM, Kutas M, Münte TF. Electrophysiological estimates of the time course of semantic and phonological encoding during listening and naming. *Neuropsychologia*. 2002; 40(7):778–787. [PubMed: 11900728]
- Sahin NT, Pinker S, Cash SS, Schomer D, Halgren E. Sequential processing of lexical, grammatical, and phonological information within Broca's area. *Science* (New York, N.Y.). 2009; 326(5951):445–449. <http://doi.org/10.1126/science.1174481>.
- Sander EK. When are speech sounds learned? *The Journal of Speech and Hearing Disorders*. 1972; 37(1):55–63. [PubMed: 5053945]

- Schmitt BM, Münte TF, Kutas M. Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. *Psychophysiology*. 2000; 37(4):473–484. [PubMed: 10934906]
- Shriberg LD. Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders. *Journal of Speech and Hearing Research*. 1993; 36(1):105–140. [PubMed: 8450654]
- Shriberg, LD.; Kent, RD. *Clinical phonetics*. Boston, MA: Pearson Education; 2013.
- Sosa AV, Stoel-Gammon C. Lexical and Phonological Effects in Early Word Production. *Journal of Speech, Language, and Hearing Research*. 2012; 55(2):596–608. [http://doi.org/10.1044/1092-4388\(2011/10-0113\)](http://doi.org/10.1044/1092-4388(2011/10-0113)).
- Strijkers K, Costa A, Thierry G. Tracking lexical access in speech production: electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex (New York, N.Y.: 1991)*. 2010; 20(4):912–928. <http://doi.org/10.1093/cercor/bhp153>.
- Szekely A, D'Amico S, Devescovi A, Federmeier K, Herron D, Iyer G, Bates E. Timed action and object naming. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*. 2005; 41(1):7–25.
- Templin, MC. *Certain language skills in children; their development and interrelationships*. Minneapolis, MN US: University of Minnesota Press; 1957.
- Turenout M, van Hagoort P, Brown CM. Electrophysiological evidence on the time course of semantic and phonological processes in speech production. *Journal of Experimental Psychology. Learning, Memory, and Cognition*. 1997; 23(4):787–806.
- Turenout M, van Hagoort P, Brown CM. Brain Activity During Speaking: From Syntax to Phonology in 40 Milliseconds. *Science*. 1998; 280(5363):572–574. <http://doi.org/10.1126/science.280.5363.572>. [PubMed: 9554845]
- Valente A, Laganaro M. Ageing effects on word production processes: an ERP topographic analysis. *Language, Cognition and Neuroscience*. 2015; 30(10):1259–1272. <http://doi.org/10.1080/23273798.2015.1059950>.
- Wellman, BI. *Physical growth and motor development and their relation to mental development in children*. Oxford England: Clark Univ. Press; 1931.
- Zhang Q, Damian MF. The time course of segment and tone encoding in Chinese spoken production: an event-related potential study. *Neuroscience*. 2009; 163(1):252–265. <http://doi.org/10.1016/j.neuroscience.2009.06.015>. [PubMed: 19524018]

Appendix. Log word frequency (LogWF), word density, and age of acquisition (AoA) data for stimuli with initial liquid and stop consonants

Word Initial Liquid				Word Initial Stop			
Word	LogWF	Density	AoA	Word	LogWF	Density	AoA
ladder	2.67	7.00	4.40	badge	2.89	12.00	6.11
ladybug	1.32	1.00	4.84	bag	3.68	26.00	4.28
lamb	2.73	23.00	4.15	ball	3.73	21.00	2.90
lamp	2.82	8.00	4.00	balloon	2.65	2.00	4.37
lasso	1.66	2.00	8.56	banana	2.74	1.00	3.78
lava	2.25	2.00	6.44	barn	2.84	10.00	4.50
lawnmower	1.40	1.00	6.11	barrel	2.73	7.00	7.72
leaf	2.42	24.00	4.60	basket	2.83	3.00	5.67
leash	2.34	15.00	5.79	bead	1.76	21.00	5.63
lemon	2.79	3.00	4.74	bear	3.47	22.00	3.58
lemonade	2.45	1.00	5.06	bed	3.98	21.00	2.89

Word Initial Liquid				Word Initial Stop			
Word	LogWF	Density	AoA	Word	LogWF	Density	AoA
letter	3.62	12.00	4.74	bee	2.72	25.00	5.00
lighthouse	2.20	1.00	6.11	beehive	1.65	1.00	7.47
lightning	2.51	0.00	4.76	bell	3.30	22.00	3.89
lime	2.23	19.00	6.61	belt	3.09	15.00	4.62
lion	2.89	2.00	4.42	bench	2.69	6.00	4.21
lipstick	2.65	1.00	5.42	bike	3.12	18.00	4.79
list	3.61	16.00	5.58	boat	3.69	27.00	3.84
lock	3.46	32.00	5.74	bone	3.12	25.00	5.53
log	2.79	12.00	6.74	book	3.96	17.00	3.68
rabbit	3.03	3.00	3.94	boots	2.99	21.00	3.89
raccoon	1.87	1.00	6.79	bottle	3.41	7.00	3.56
racquet	1.26	4.00	8.20	bowl	3.04	19.00	4.26
rain	3.40	28.00	3.60	box	3.66	3.00	4.30
rainbow	2.61	1.00	4.26	bus	3.58	17.00	3.85
raisins	2.05	1.00	5.05	butterfly	2.45	1.00	3.67
rake	2.18	26.00	5.32	button	3.16	5.00	4.78
rat	3.22	31.00	5.11	deer	2.65	27.00	5.17
rattle	2.24	7.00	4.50	desk	3.35	2.00	5.56
rectangle	1.08	1.00	5.00	detective	3.49	1.00	6.58
rhino	2.03	1.00	6.00	dice	2.73	16.00	6.37
ribbon	2.41	2.00	5.25	dinosaur	2.31	1.00	4.91
rice	2.89	18.00	3.72	dock	2.70	21.00	8.22
ring	3.67	21.00	4.53	dog	3.99	8.00	2.80
robe	2.64	15.00	6.38	doghouse	1.76	0.00	6.74
robot	2.79	1.00	5.35	doll	3.10	14.00	3.68
rock	3.64	20.00	3.22	dolphin	2.15	1.00	6.05
rocket	2.78	5.00	5.63	donkey	2.44	1.00	6.00
rope	3.06	23.00	5.44	dot	2.53	24.00	4.33
rose	3.43	22.00	6.11	doughnut	2.33	1.00	5.42
				dove	2.45	14.00	6.63
				duck	3.10	22.00	3.50
Mean	2.58	10.33	5.31	Mean	2.95	12.57	4.87
Median	2.65	6.00	5.18	Median	2.94	14.00	4.56

HIGHLIGHTS

- ERP responses to words with initial liquid and stop consonants differed between 293 ms and 371 ms.
- Words beginning with liquid consonants elicited larger N2 responses.
- Data indicate late onset of phonological code retrieval, consistent with previous picture naming studies.
- Neural evidence was found in support of the notion that phonemes are distinguished by their distinctive features and phonological complexity.

Figure 1.

ERP waveforms for the 42 electrodes from which amplitude measures were made.

Figure 2.
View of three mid-line electrodes (FCz, Cz, CPz) showing P2-N2-P3 peaks in the ERP waveform.

Figure 3.

Raster diagram illustrating differences between ERPs elicited by object labels beginning with liquids and stops. Black rectangles indicate electrodes/time points in which the ERPs for liquids are more negative than those for stops. The color scale indicates the size of the t-test result: The darker the color, the higher the level of significance. Note that the electrodes are organized along the y-axis somewhat topographically. Electrodes on the left and right sides of the head are grouped on the figure's top and bottom, respectively; midline electrodes are shown in the middle. Within those three groupings, y-axis top-to-bottom corresponds to scalp anterior-to-posterior. One outcome was apparent: a broadly distributed effect from 293–371 ms, signifying the time period during which object labels beginning with liquids elicited larger N2 responses than did those with initial stops. The smallest significant t-score was: $t(33) = -2.037$, $p < .03$, which had a test-wise alpha level of 0.0497.

Figure 4.

Mid-line electrode responses to words with initial stops and liquids that had high (A) and low (B) frequencies. In the high frequency analysis, cluster-based permutation tests revealed no significant consonant class effect between 150 and 400 ms, although visual inspection indicated that words with initial liquids elicited larger N2 responses than did stops. By contrast, significant and extended consonant class effects for the low frequency words beginning with liquids were indicated. These words elicited significantly larger N2 responses across nearly the entire 250 ms analysis window. Responses to words with initial liquid and stop consonants were also combined into the third set of waveforms to examine overall frequency differences (C). The ERP waveforms revealed no statistical differences between high and low frequency words.

Figure 5.

Raster diagram illustrating differences between ERPs elicited by low frequency words with initial liquids and stops. Two broadly distributed significant clusters were identified from 187 to 254 ms and from 280 to 400 ms. During these periods, words beginning with liquids elicited larger N2 responses than did their counterparts with initial stops. The smallest significant t-score was: $t(33) = -2.035$, $p < .02$, which had a test-wise alpha level of 0.0499.

Figure 6.

Mid-line electrode responses to one-syllable and multisyllabic words beginning with (A) stop and (B) liquid consonants. Some variation was observed in the ERP waveform for liquid and stop consonants. Cluster-based permutation tests indicated no significant syllable-length effects during the proposed phonological processing time windows (150–400 ms). Due to the small number of trials, responses to one-syllable and multisyllabic words with initial liquid and stop consonants were also combined into the third set of waveforms (C). The combined ERP waveforms had no statistical differences.

Figure 7.

Experimental procedure. Participants were presented single color cartoon pictures, one at a time, on a computer screen. They were asked to overtly name the objects as quickly as possible. Verbal responses were recorded by a voice-activated microphone that also recorded naming response time.

Figure 8.
Nine electrode groups selected by scalp location based on Strijkers et al. (2010).

Table 1

Accuracy and verbal response time data for subject and item analyses

Analysis Type	Word Initial Consonant	Percent Accurate	Verbal Response Time (ms)
Subject	Stop	91.9 (0.7)	1126 (27)
	Liquid	90.7 (1.1)	1155 (26)
Item	Stop	89.4 (2.5)	1204 (21)
	Liquid	91.0 (2.6)	1248 (22)

Note: Numbers in parentheses represent standard error of the mean.

Table 2

Mean ERP amplitudes (in microvolts) in the nine electrode groups for P2 and N2 peaks

P2 Mean Amplitude in the 150–275 ms post-picture onset time window										
Stimulus Type	Left Frontal	Fronto-Central	Right Frontal	Left Central	Centro-Parietal	Right Central	Left Parietal	Occipital	Right Parietal	Overall
Stop	−2.464 (.680)	−2.486 (.738)	−2.597 (.672)	−1.376 (.640)	−0.855 (.691)	−1.363 (.588)	0.957 (.518)	3.113 (.496)	1.280 (.466)	−0.644 (.507)
Liquid	−3.098 (.703)	−3.039 (.751)	−3.075 (.697)	−2.007 (.651)	−1.407 (.694)	−1.923 (.607)	0.622 (.570)	2.885 (.637)	0.890 (.521)	−1.128 (.548)
Overall	−2.781 (.672)	−2.763 (.724)	−2.836 (.662)	−1.692 (.628)	−1.131 (.668)	−1.643 (.576)	0.790 (.512)	2.999 (.540)	1.085 (.463)	
N2 Mean Amplitude in the 275–400 ms post-picture onset time window										
Stimulus Type	Left Frontal	Fronto-Central	Right Frontal	Left Central	Centro-Parietal	Right Central	Left Parietal	Occipital	Right Parietal	Overall
Stop	−2.827 (.790)	−3.071 (.853)	−3.060 (.695)	−0.481 (.732)	0.660 (.748)	−0.415 (.591)	3.075 (.636)	4.812 (.676)	3.580 (.570)	0.253 (.577)
Liquid	−3.317 (.766)	−3.702 (.818)	−3.669 (.687)	−1.262 (.739)	−0.202 (.817)	−1.361 (.618)	2.329 (.712)	4.292 (.819)	2.720 (.652)	−0.464 (.626)
Overall	−3.072 (.759)	−3.387 (.814)	−3.365 (.667)	−0.871 (.712)	0.229 (.752)	−0.888 (.574)	2.702 (.646)	4.552 (.720)	3.150 (.577)	

Note: Numbers in parentheses represent standard error of the mean.

Table 3

Stimulus characteristics

	Word Initial Stop	Word Initial Liquid
Number of 1 syllable words	28	17
Number of 2 syllable words	10	19
Number of 3 syllable words	4	4
Mean log-frequency	2.95 (0.60)	2.58 (0.67)
Mean age of acquisition (years)	4.87 (1.33)	5.31 (1.14)
Mean neighborhood density	12.57 (9.50)	10.33 (10.21)

Note: Numbers in parentheses following means are standard deviations.