

Planning and Articulation in Incremental Word Production: Syllable-Frequency Effects in English

Joana Cholin

Basque Center on Cognition, Brain and Language (BCBL),
Donostia, Spain

Gary S. Dell

Beckman Institute, University of Illinois at Urbana-Champaign

Willem J. M. Levelt

Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands

We investigated the role of syllables during speech planning in English by measuring syllable-frequency effects. So far, syllable-frequency effects in English have not been reported. English has poorly defined syllable boundaries, and thus the syllable might not function as a prominent unit in English speech production. Speakers produced either monosyllabic (Experiment 1) or disyllabic (Experiment 2–4) pseudowords as quickly as possible in response to symbolic cues. Monosyllabic targets consisted of either high- or low-frequency syllables, whereas disyllabic items contained either a 1st or 2nd syllable that was frequency-manipulated. Significant syllable-frequency effects were found in all experiments. Whereas previous findings for disyllables in Dutch and Spanish—languages with relatively clear syllable boundaries—showed effects of a frequency manipulation on 1st but not 2nd syllables, in our study English speakers were sensitive to the frequency of both syllables. We interpret this sensitivity as an indication that the production of English has more extensive planning scopes at the interface of phonetic encoding and articulation.

Keywords: language production, word-form encoding, syllable-frequency effects, speech planning, articulation

“Do Speakers Have Access to a Mental Syllabary?” is the title of a highly regarded and often-cited article by Levelt and Wheeldon (1994) that asked whether syllables are separately stored units that are retrieved from memory during speech planning. The idea of a store that contains abstract motor programs (articulatory routines) of syllabic size goes back to Crompton (1981). Crompton proposed that speakers use precompiled syllable programs rather than computing these programs every time anew when needed for articulation. He argued that precompilation would facilitate rapid speech production and that the syllable was the unit best suited to such precompilation (see also MacNeilage, 1998, for arguments for such a unit on evolutionary grounds).

Levelt and Wheeldon (1994) were the first who attempted to identify stored syllabic units. They did so by looking for syllable-frequency effects in speech production, applying the logic that frequency effects are expected only for stored units (Jescheniak & Levelt, 1994; Oldfield & Wingfield, 1965). Specifically, they investigated naming latencies for two-syllable words consisting of high- versus low-frequency syllables. Levelt and Wheeldon found faster naming latencies for words consisting of high-frequency syllables compared with words containing low-frequency syllables when word frequency was controlled for. The obtained syllable-frequency effects therefore support the notion of a mental syllabary where syllables are stored separately from words (see also Aichert & Ziegler, 2004; Carreiras & Perea, 2004; Cholin, Levelt, & Schiller, 2006; Laganaro & Alario, 2006).

Furthermore, Levelt and Wheeldon (1994) specifically found an effect of the second syllable’s frequency, reporting that the “bulk of the syllable-frequency effect is due to the word-final syllable” (Levelt & Wheeldon, 1994, p. 260). They concluded that the syllables of a word may be retrieved from a frequency-sensitive mental syllabary and that speakers start articulation only when all of those syllables are ready to be executed. The final syllable of a word will often be the last to be ready, and hence its properties will have a large influence on production times.

Recent findings, however, have questioned the conclusion that syllable-frequency effects in production inhabit the final syllable of two-syllable words. Studies that systematically manipulated the first and the second syllable (of disyllabic pseudowords) in Dutch and Spanish found frequency effects for only the first syllable (Carreiras & Perea, 2004; Cholin et al., 2006). The argument that was put forward in Cholin et al. (2006) and that is in line with the

Joana Cholin, Basque Center on Cognition, Brain and Language, Donostia, Spain; Gary S. Dell, Beckman Institute, University of Illinois at Urbana-Champaign; Willem J. M. Levelt, Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands.

Joana Cholin would like to express her gratitude for support given by the Netherlands Organization for Scientific Research (NWO; Grant S 29-1) and by the Spanish Science and Innovation Ministry (MICINN; Grant RYC200903917). The laboratory used for Experiments 1–3 was supported by U.S. National Institutes of Health Grants DC000191 and HD44458. The authors wish to thank Susan Garnsey and Sara Finley for recording the acoustic stimuli, John Nagengast and Johan Weustink for their technical support, Keren Shatzman for her help with material selection, Delphine Dahan for making it possible to run Experiment 4 in her lab, and Sarah Johnstone Drucker for her organizational help.

Correspondence concerning this article should be addressed to Joana Cholin, Basque Center on Cognition, Brain and Language (BCBL), Paseo Mikeletegi 69, 20009 Donostia-San Sebastián, Spain. E-mail: j.cholin@bcbleu

findings and conclusions of many other studies (see e.g., Meyer, Roelofs, & Levelt, 2003) was that speakers can start articulation with a syllable that has been completed even though the second syllable is still under construction, thereby not showing effects of the frequency of the second syllable on word initiation. One difference between the studies that found effects for the first but not for the second syllable (Carreiras & Perea, 2004; Cholin et al., 2006) and the Levelt and Wheeldon (1994) study was that the former used pseudowords whereas the latter used existing words. Waiting to speak until all syllables are phonetically encoded makes sense for words as opposed to pseudowords. Consider the analogy of arriving in a foreign country. Related persons (analogous to the syllables of a familiar word) move through the various immigration/custom steps together, each person waiting for their relative to finish a step before jointly moving on to the next one. People traveling alone (the syllables of a pseudoword) proceed through the steps as individuals and ignore the progress of those behind them in the queue. There are, however, other studies testing *words* that support the claim that speakers (can) start articulation upon phonetic completion of the first syllable (see e.g., Bachoud-Lévi, Dupoux, Cohen, & Mehler, 1998; Meyer et al., 2003; Schriefers & Teruel, 1999; see also Meyer, Belke, Häcker, & Mortensen, 2007). In terms of the analogy just described, this would mean that under certain circumstances even members of a family proceed through immigration without waiting for each other.

The present study investigated syllable-frequency effects in mono- and disyllabic pseudowords in English spoken production. So far, the languages in which significant syllable-frequency effects in production tasks have been found are Dutch (Cholin et al., 2006; Levelt & Wheeldon, 1994), German (Aichert & Ziegler, 2004), Spanish (Carreiras & Perea, 2004), and French (Laganaro & Alario, 2006)—that is, languages that have comparatively clear syllable boundaries and, thus, in which (stored) syllabic units might play a more prominent role (see also Chen, Chen, & Dell, 2002, for evidence for stored syllabic units in Mandarin Chinese). English, on the other hand, has less obvious syllable boundaries, and syllables may therefore not constitute separately stored production units. Native speakers of English have inconsistent intuitions about where syllables begin and end (see e.g., Treiman & Danis, 1988; Treiman & Zukowski, 1990). This may be due to the high degree of ambisyllabicity in English (see e.g., Kahn, 1980; Selkirk, 1982; Spencer, 1996). Ambisyllabic consonants constitute the coda of one syllable as well as the onset of the following syllable (e.g., in the word “*happy*”), thus making it unclear about where the syllable boundary lies (e.g., before or after the /p/). Native English speakers know that “*happy*” consists of two syllables, but they disagree as to whether the /p/ belongs to the first or second syllable or both (see e.g., Treiman & Danis, 1988). Linguistic theory (see e.g., Pulgram, 1970) as well as psycholinguistic experimentation (see e.g., Bradley, Sánchez-Casas, & García-Albea, 1993; Cutler, 1995; Cutler, Mehler, Norris, & Segui, 1986; Cutler, Norris, & Williams, 1987) converge on the conclusion that syllable boundaries in English words are not clear-cut. Moreover, these English studies contrast with studies using languages such as French and Spanish, which found that the syllable plays a prominent role in the segmentation of spoken words (see e.g., Dupoux, 1993; Mehler, Dommergues, & Frauenfelder, 1981) as well as in the recognition of visual stimuli (see e.g., Álvarez, Carreiras, &

Perea, 2004; Taft, Álvarez, & Carreiras, 2007). It has been proposed that, instead of syllables, the *rime* (i.e., the vowel and coda of a syllable, such as the /at/ in /cat/ or /rat/) serves as a crucial phonological/phonetic unit in English (see e.g., Dell, Juliano, & Govindjee, 1993; Lee & Goldrick, 2008; Treiman, 1983; Treiman & Danis, 1988). For example, Lee and Goldrick (2008) tested English speakers’ awareness of subsyllabic representations and found a sensitivity to the (frequency of the) co-occurrence of vowel–coda sequences. So far, frequency effects for syllabic units in English have not been found (Croot & Rastle, 2004), and thus it remains an open question whether the syllable is a stored phonetic unit in production generally (i.e., also in English). Furthermore, the investigation of disyllables in English might offer specific insights into planning processes in languages with vague syllable structures.

The Process of Speech Planning and Syllabary Access

When speakers produce sentences or phrases, the speech-planning system starts by generating a conceptual message that provides the underlying structure as well as the semantic content. The subsequent encoding steps involve the lexical retrieval of word units that comply with the structure and the content of this message. Most models distinguish between the retrieval of word units, the so-called lemmas, and the retrieval of their associated word forms (see e.g., Dell, 1986, 1988; Levelt, 1989; Levelt, Roelofs, & Meyer, 1999). Whereas serial models assume that lemmas are selected before their corresponding word forms are activated (see e.g., Levelt et al., 1999), interactive two-step models (Dell, 1986, 1988; see also Dell, Schwartz, Martin, Saffran, & Gagnon, 1997) assume that lemmas and word forms are activated in parallel. The activated word form will lead to an activation of the morpheme(s) that a word form consists of, which will, in turn, trigger the spell-out of the morphemes’ single segments (i.e., rose [+plural] will activate /r/ /o/ /v/ /z/ + /(1)z/). Whether this string of segments is connected to syllabic slots within a metrical frame and is specified for syllable-internal positions (such as onset or rime) as assumed by Dell (1986, 1988) or whether the segments are not yet assigned to any subsyllabic positions but only have numbered links to indicate their serial order within their corresponding word forms as assumed by Levelt et al. (1999; see also Roelofs, 1997a) is still subject to debate (for an overview see Schiller & Costa, 2006). The Levelt et al. model assumes that the segmental string will then be syllabified and assigned stress according to constraints that respect the phonological properties of the language (Schiller, Fikkert, & Levelt, 2004; Schiller, Meyer, Baayen, & Levelt, 1996). The end product of phonological encoding is a phonological word consisting of abstract phonological syllables. Thus, up to this point in the process, the model assumes there are no stored syllable units. Stored syllable units become important in the next step, phonetic encoding.

During phonetic encoding, these abstract phonological syllables will be converted into phonetic syllables. This conversion can be done by translating the single segments of a syllable one by one, which is referred to as the *online assembly of segments*, or by accessing the already assembled syllables that are stored in the mental syllabary. Retrieving precompiled syllables (or abstract motor programs thereof) should be less effortful than the online computation of syllables and, as already pointed out, sensitive to

effects of syllable frequency. The stored units are packages of abstract motor programs of single segments bundled together to form entire syllables. The more frequently a given phonological environment (re-)occurs, the more likely it is that the resulting bundling will be a stored unit within the mental syllabary. In this sense, it is not only the canonical lexical entries within the mental lexicon but also the recurring phonological contexts that can yield high-frequency syllables. For example, the syllable /zɪz/ is a high-frequency English syllable that emerges when plural or third-person singular forms of words ending in /z/ (e.g., *ro.ses*, *a.ccu.ses*) are generated. Syllables like /zɪz/ are not necessarily part of the lexicalized word forms but surface (only) in certain inflectional forms. Moreover, in very fast, connected speech, new syllables may arise that are the result of assimilations; for example, phonemes or entire syllables may be dropped so that new phonological environments evolve that can lead to noncanonical syllables (for a discussion of the phenomenon of *resyllabification*, see Schiller et al., 1996). If these syllables arise frequently, their storage within the syllabary is likely.

In short, models of word-form production, and specifically the model of Levelt et al. (1999), propose that stored syllable units exist at a postlexical, or phonetic, level. Syllable-frequency effects in production latencies that are independent of word frequency constitute the best existing support for the general notion that syllabic units are stored (see e.g., Cholin et al., 2006; Levelt & Wheeldon, 1994). Moreover, recent studies have suggested that such effects are, as hypothesized, located at the phonetic level, that is, at the level of the hypothesized syllabary (Cholin & Levelt, 2009; Laganaro & Alario, 2006).

Retrieving stored representations within the hypothesized syllabary during production is much like the process of retrieving a higher order unit from recognized phonological units during word recognition. First, potential syllable programs are activated, followed by the selection of the one (or more in the case of multisyllabic words) that best matches the syllabified segment string (Roelofs, 1997a, 1997b). A stored syllable within the syllabary becomes active as soon as a part of it (e.g., its onset) has been generated during phonological encoding. Hence, not only the eventual target syllable but also syllable programs that only partly match the output of phonological encoding become active as well. Thus, there is always an activation of multiple syllables within the syllabary, (a) because of the activation of all stored syllables that exhibit a segmental overlap with the addressing phonological syllable and (b) because the activation of syllables triggered by the second and potential subsequent syllables of a phonological word might occur before the selection of the first syllable is completed. Because of the assumed activation of multiple syllables when the syllabary is accessed, one can conceive of the notion of an active syllable “neighborhood,” analogous to the lexical-phonological neighborhoods that influence word recognition (see e.g., Carreiras, Álvarez, & de Vega, 1993; Carreiras & Perea, 2002) and production (see e.g., Baus, Costa, & Carreiras, 2008; Vitevitch, 1997, 2002; Vitevitch & Sommers, 2003). It is unclear, though, whether increases in syllable neighborhood density would speed or retard retrieval of a syllable program. Dense lexical neighborhoods tend to facilitate production (see e.g., Gordon & Dell, 2001; Vitevitch, 2002) but slow down spoken-word recognition (see e.g., Vitevitch, Luce, Pisoni, & Auer, 1999). The retrieval of a syllable program from segmental input is akin to both production (because it is

production) and recognition, because it involves retrieval of a unit (a syllable) from unit parts (segments) whose distribution defines the neighborhood. In either event, it is important for our experimental manipulations of syllable frequency to minimize any potential confounding between frequency and neighbor distribution. The particular method of choosing syllables that we adopt does so with regard to one plausible notion of a syllable’s neighborhood.

Planning Scope and the Initiation of Articulation

Word-form encoding is a serial process that gives the first syllable a head start in accessing the syllabary (see e.g., Meyer, 1990, 1991; Wheeldon & Levelt, 1995). This temporal advantage allows the system to complete encoding of the first syllable before any other, consecutive syllable is selected. Whether this first syllable will be uttered as soon as the articulators get hold of it (like a single traveler proceeding directly through the immigration/custom steps) or whether the first syllable is kept in an articulatory buffer until all syllables of the multisyllabic word are ready to be executed (the traveler in a group does not move to the next step until her group is processed) is part of our investigation.

The moment in time that articulation is initiated seems to depend not only on the availability of the first encoded unit but also on the respective *planning scope*, that is, the extent to which encoding must lead articulation. The scope of planning can span more or fewer elements or linguistic units. The theoretical positions and empirical findings on this question support both fixed-size and flexible planning scopes (for a discussion see Alario, Costa, & Caramazza, 2002a, 2002b; Levelt, 2002; see also Damian & Dumay, 2007). A fixed scope would demand that a predetermined set of linguistic elements be encoded before speakers start overt articulation (which also implies that a speaker must start articulation as soon as encoding for this set is completed). A flexible planning scope, on the other hand, is variable, and the assessment of what unit is most appropriate can vary between linguistic contexts, communicative situation, and speakers (see Schriefers & Teruel, 1999).

Here, we are asking whether the syllable is the unit that controls the delay between phonetic encoding and articulation. The research on Dutch and Spanish, showing frequency effects for frequency-manipulated first but not second syllables, suggests that as soon as the first syllable of a disyllabic word is phonetically encoded, articulation can be initiated. The experiments reported here investigate whether English also uses a syllable-sized planning scope or whether a planning scope of a different size is applied.

Overview of the Present Study

Experiment 1 tests the production of high- and low-frequency monosyllabic pseudowords. If syllables are stored entities that are accessed during English speech production, we expect to find that high-frequency syllables are produced faster than low-frequency syllables. Experiments 2–4 then investigate syllable-frequency effects in disyllabic pseudowords with high- and low-frequency first (Experiment 2) or second (Experiments 3 and 4) syllables to test which planning scope English speakers use for articulation. In all experiments, a symbol–association learning task (Cholin et al., 2006) is used to contrast the production latencies for high- and low-frequency syllables. The participants’ task is to respond to a

visually presented production cue by producing a specific syllable (or disyllable) as quickly as possible. There were two possible cues, distinguished by whether the cue was presented to the left or right side of the computer screen. These left and right positional cues had previously been associated with the respective target syllables (or disyllables) of one pair of items. Each cue thus told the participant which member of the pair to produce on that trial. A detailed description of the paradigm is provided next.

The Basic Material

All frequency counts were obtained from the computer database CELEX (CEntre for LEXical Information; Baayen, Piepenbrock, & Gulikers, 1995), which has an English lexicon based on 17.9 million word tokens. Syllable frequency was counted for phonetic syllables in English.¹ The phonetic script differentiates the reduced vowel schwa from full vowel forms, giving 12,262 different syllable types. Syllable frequencies were calculated for the database from the word form occurrences per one million. Our principal measure of frequency is the *summed frequency of occurrence* of each syllable in the corpus. For each syllable, the sum includes all instances of the syllable, regardless of its position in the word.² This measure ranged between 0 and 61,092 occurrences per one million words. We also took note, as an additional check on our syllables, of how many lexical items each syllable occurred in for the corpus, that is, a frequency count through the lexicon rather than through the corpus. This *number of occurrences in the lexicon* count was examined to ensure that the summed counts came from multiple words, particularly in the case of high-frequency syllables.

The experimental high- and low-frequency items should differ only in their syllable frequency. Therefore, it was crucial to construct an experimental item set that was controlled for number of phonemes, phoneme frequency, consonant–vowel (CV) structure, bigram/biphoneme frequency, and neighborhood density (i.e., basically everything except for syllable frequency).³ To control for all of these factors, we limited the studies to CVC syllables that were constructed in a specific manner. Sets of four syllables were paired into one syllabic quartet: The same syllable body (e.g., /zɪ/) appeared once in a high-frequency syllable /zɪz/ as well as in a low-frequency syllable, /zɪn/. The rimes of those syllables (/ɪz/ and /ɪn/) also appeared in a high- and in a low-frequency syllable, although in the contralateral way: in /gɪn/ (high frequency) and in /gɪz/ (low frequency; see Table 1 for a schematic depiction of one quartet).

Due to this specific matching, all potential confounds arising from specific segments or binary segmental combinations were eliminated. Also, note that each matched high- and low-frequency pair has exactly the same syllabic neighbors, where a *syllabic*

neighbor is defined as a syllable that differs from a target by only one segment in a corresponding position and where the shared segments are contiguous. This is arguably the most relevant definition for neighbors at a phonetic/articulatory level, where segmental position and coarticulation are potentially important. In total, eight of those syllabic quartets were constructed to serve as materials, resulting in 16 high- and 16 low-frequency syllables. The 16 high-frequency items had a summed frequency of occurrence ranging from 5.98 to 774.24 per million words, with an average of 201.34 ($SD = 225.92$). The mean number of occurrences in the lexicon per syllable for these items was 163 ($SD = 158$). Low-frequency syllables ranged from zero to 8.10 for summed frequency, with an average of 3.39 ($SD = 2.68$) and 13 ($SD = 7$) for number of occurrences in the lexicon. Table A1 in Appendix A contains a full list of experimental items and their frequency values. Although care was taken that none of the syllables are existing words in English, one quartet consisted of syllables that are also proper names in English: Jen, Bess, Ben, and Jess (Quartet No. 8). Given the limited number of syllable quartets that could be found in English, we decided to include this quartet nevertheless.

Acoustic versions of the syllables were spoken by a female native speaker of American English. The spoken syllables were digitized at a sampling rate of 22 kHz, to be used during the learning phase of the experiment. They varied in duration from 472 ms to 728 ms, with an average of 596.59 ms ($SD = 69.72$). There was no difference in duration between high-frequency and low-frequency syllables (both $ts < 1$).

Monosyllabic Pseudowords (Experiment 1)

Method

Participants. Thirty-two participants from the psychology participants pool of the University of Illinois at Urbana–Champaign took part in the experiment. All of them were undergraduate students and participated in exchange for course credit.

¹ Note that the phonetic syllable corpus was derived from a morphological corpus that included inflections and derivations such as *ro.ses* and *pro.tec.ted*. Dots indicate syllable boundaries.

² Due to the exacting search criteria for the syllabic material, the positional frequency of a given syllable (i.e., the frequency of syllables stemming from specific positions within words) could not be taken into account. However, former experiments using the same material construction and the same paradigm as in the present study showed that high-frequency syllables that occur only as final syllables in existing Dutch words yielded significant syllable-frequency effects when presented as first syllables in disyllabic (Dutch) pseudowords (Cholin et al., 2006). This finding supports the theoretical argument that the stored syllables within the syllabary are not specified for their position within words, though there might be scenarios in which one syllable might in fact be stored as “ba-first syllable” versus “ba-last syllable.”

³ The importance of controlling for segmental frequency became apparent in a follow-up study of the original one by Levelt and Wheeldon (1994). In an attempt to replicate those results, Levelt and Meyer (reported in Hendriks & McQueen, 1996) ran an experiment in which a large number of possible confounding factors were controlled for, and neither syllable nor segment-frequency effects were obtained.

Table 1
An Experimental Quartet Consisting of Two High- and Two Low-Frequency Syllables

Syllables within one quartet	
High frequency	Low frequency
zɪz	zɪn
gɪn	gɪz

Note. Onsets and offsets are frequency-controlled.

They had no known hearing deficit, and they had normal or corrected-to-normal vision.

Experimental design. The two-level variable frequency (high vs. low) was tested within participants and within quartets. Each participant produced each of the 32 syllables eight times, resulting in a total of 256 experimental trials. The 32 syllables were divided into sets of two. These sets were frequency-homogeneous (i.e., a high-frequency syllable of one quartet was combined with a high-frequency syllable of another quartet), and the same was done with their low-frequency counterparts (e.g., /fæk/, high frequency – /nɪʃ/, high frequency; /fæl/, low frequency – /nɪd/, low frequency). The pairing of two syllables resulted in 16 experimental sets, eight high- and eight low-frequency sets. There was no segmental overlap of syllables within sets or between consecutive sets. Moreover, high- and low-frequency counterparts (e.g., /fæk/ – /nɪʃ/; /fæl/ – /nɪd/) were presented with the maximum number of eight sets between them. High- and low-frequency item sets alternated across the 16 item sets. The sequence of items was controlled by a Latin square. Thirty-two experimental versions resulted: Every item set occurred at each of the 16 possible positions across all experimental versions. The presentation of the individual syllables at the two possible positions (left vs. right) was also counterbalanced across participants. See Table A2 in Appendix A for the pairing of syllables into frequency-homogeneous sets.

In order to prevent the direct repetition of identical items in the present two-item design, we introduced fillers in the form of different numbers presented in the center of the computer screen that had to be named in between experimental trials. This number naming should distract participants' attentiveness to any expectation of trial succession. Moreover, it should help to neutralize the articulators because the immediate repetition of two identical items could have huge facilitation effects. Four monosyllabic numbers (1, 2, 8, and 9) were selected as fillers. There was no phonological overlap between these numbers and the experimental items. Fillers and experimental trials were presented in alternating order, with a filler trial always serving as the first item in each experimental block.

Procedure. The experiment consisted of alternating learning, practice, and test phases. Participants were tested individually in a quiet room.

In each learning phase, the participant learned to associate the two syllables of a stimulus pair (e.g., /fæk/ – /nɪʃ/; see Table A2 in Appendix A) with the left or the right positions on the computer screen (e.g., /fæk/ with the left side and /nɪʃ/ with the right side). An icon of a little white loudspeaker (4 by 4 cm) was presented on the right or the left side of a black computer screen while simultaneously the to-be-learned target presented auditorily via headphones. Participants were asked to memorize the side associated with each target syllable. Each target syllable was presented twice together with the symbol in its specific position.

In the practice phase, the symbols were simultaneously presented on both sides and one of the target items was presented via headphones. Participants then had to demonstrate that they remembered the association between screen side and the syllables by pressing a spatially congruent button on a dual box. Each target was presented four times. Only if participants passed the respective practice phase without errors was the test phase started; otherwise, the learning and practice phase were repeated. Participants were

explicitly asked to refrain from rehearsing the target items by articulating them.

In the test phase, filler trials (i.e., number naming) and experimental trials alternated. Each test phase started with a filler trial. In each experimental trial, the loudspeaker symbol was presented in one of the two screen positions to cue target production (e.g., if it appears on the left, the participant should produce /fæk/). In the filler trial, one of the four numbers (1, 2, 8, or 9) was presented in the middle of the computer screen and had to be named. Simultaneously with cue presentation, the voice key was activated for 1,500 ms. The symbol disappeared after the response with a delay of 500 ms. Then 100 ms later, the next (filler) trial started. Each target had to be produced eight times in a random order. Thus, each experimental block consisted of eight repetitions of each of the two members of an experimental pair (e.g., eight times /fæk/ and eight times /nɪʃ/, alternating with 16 number trials). The experiment started with two practice sets consisting of practice items mimicking the experimental materials. The duration of the experiment was approximately 50 min.

The presentation of the stimuli and the measuring of the reaction times (RTs) were controlled by the NESU (Nijmegen Experiment Setup Unit) software package developed at the Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands. The spoken reactions were registered by a microphone, which fed into a NESU box voice key device and a digital recorder. The experimenter sat in the same room as the participant and took note of hesitations, voice key errors, wrong naming responses, and time-outs.

Results

Test items leading to wrong or invalid responses (mispronunciations, voice key errors, and hesitations) were coded as errors and were not included in the RT analysis. RTs above 1,500 ms and below 200 ms were also considered to be invalid responses and did not enter the RT analysis. Observations deviating from a participant's and an item's mean by more than two standard deviations were considered as outliers and were also discarded from the RT analysis. In total, 359 (4.4%) trials were treated as errors and 161 (2.0%) as outliers.

The mean RTs and error rates for the high- and low-frequency items were submitted to *t* tests. Two complementary analyses were computed, one treating participants (t_1) and one treating quartets (t_2) as random factors. The mean voice onset latencies, standard errors, and error rates for Experiment 1 are summarized in Table 2.

The analysis of RTs showed that high-frequency monosyllabic pseudowords were produced significantly faster than were low-frequency monosyllabic pseudowords, $t_1(31) = 2.36$, $p = .0247$, $\eta_p^2 = .15$; $t_2(7) = 2.95$, $p = .0214$, $\eta_p^2 = .55$. The analysis of errors

Table 2
Mean Voice Onset Latencies (in Milliseconds) and Percentage of Errors for Experiment 1

Frequency	<i>M</i> (<i>SE</i>)	<i>Errors</i> (<i>SE</i>)
High	464 (8)	4.7 (0.5)
Low	469 (8)	4.0 (0.5)
Difference scores	–5 (2)	0.7 (0.6)

yielded no significant effects, $t_1(31) = 1.14$, $p = .2630$, $\eta_p^2 = .04$; $t_2(7) = 1.06$, $p = .3243$, $\eta_p^2 = .14$.

Discussion

High-frequency syllables were associated with shorter naming latencies than were low-frequency ones. The effect was small (5 ms) but clearly reliable. Because all others factors (e.g., onset phonemes, length, phoneme frequency, bigram/biphone frequency, neighborhood) have been equated by the use of syllable quartets, this result strongly suggests that entire syllables in fact constitute stored entities in English. The result is particularly noteworthy given that English syllable boundaries are less clear-cut. To our knowledge, these data are the first to show significant syllable-frequency effects in English. A previous study by Croot and Rastle (2004) also found a nonsignificant 5-ms difference between the production of English high-frequency and nonexistent but phonotactically legal English syllables. So, Croot and Rastle's result is numerically similar to ours. Because our study had eight repetitions per syllable and manipulated frequency level within matched syllable quartets, it had the power to detect effects of this size. Notice that the syllable-frequency effects obtained in other languages were also small (i.e., Dutch: 9–10 ms and 12–19 ms, Cholin et al., 2006, and Levelt & Wheeldon, 1994, respectively; Spanish: 10–14 ms, Carreiras & Perea, 2004).

The next experiments, investigating syllable-frequency effects in disyllabic pseudowords, aim to shed light on the question of whether the first and/or the second syllable will be sensitive to frequency effects. The relative lack of clear syllables in English makes it conceivable that the planning scope, and hence the initiation of articulation, might be different compared to languages in which syllable boundaries are clearer. English's less-transparent boundaries might lead to less chunking into discrete syllables, hence favoring a larger planning scope. In other words, an effect of the first syllable is expected, and an effect of the second syllable seems possible.

Disyllabic Pseudowords With the Frequency Manipulation on the First Syllable (Experiment 2) and the Second Syllable (Experiment 3)

The next two experiments investigate high- and low-frequency first and second syllables in disyllabic pseudowords to test whether (a) syllable-frequency effects will also be found in this context and (b) whether we will find an effect for the first and/or for the second syllable in English.

Method

Participants. In each of the following two experiments, 32 participants drawn from the same population as in Experiment 1 were tested. None of them took part in more than one experiment.

Materials. In order to construct the disyllabic pseudowords, we took the CVC syllables that revealed significant syllable-frequency effects in Experiment 1 and combined them with high-frequency CV syllables. One high-frequency syllable was assigned to one quartet to serve as the first syllable (Experiment 2) or second syllable (Experiment 3) for all four members of this quartet (see Table 3).

By always appending the same CV syllable to all four members of one quartet, we controlled for the transition of the first to the second syllable within the disyllabic pseudowords; that is, the frequency of the bigrams/biphones is the same for high- and low-frequency syllables within one quartet. The list of the eight high-frequency syllables can be found in Table B1 of Appendix B. The CV syllables ranged in summed frequency from 471.01 to 25,401.01, with an average of 5,998.47 ($SD = 8,164.95$) per one million words; the mean number of occurrences in the lexicon per syllable for these items was 1,241 ($SD = 702$). We opted for high-frequency appended syllables (among the first percentile of the most high-frequency English syllables) because these syllables are most likely to be stored within the syllabary. The retrieval of those high-frequency syllables should therefore be fast and least error-prone. Care was taken that none of the resulting disyllabic pseudowords would form any existing word in English. The grouping of two items into one experimental set was the same as the original pairing in Experiment 1 (see Table B2 of Appendix B for the materials used in Experiments 2 and 3). The same numerals that were named in Experiment 1's filler trials were also included in Experiments 2 and 3 to separate critical disyllable production trials.

The spoken disyllabic pseudowords were digitized at a sampling rate of 22 kHz. High- and low-frequency items used in Experiment 2 varied in duration from 558 ms to 812 ms, with an average of 685 ms ($SD = 69$). There was no difference in duration between high- and low-frequency syllables, $t(30) = 1.18$, $p = .25$. High- and low-frequency items in Experiment 3 varied in duration from 558 ms to 872 ms, with an average of 684 ms ($SD = 82.97$). There was no difference in duration between high- and low-frequency syllables, $t(30) = 1.62$, $p = .12$.

The disyllabic pseudowords were spoken by the same speaker as in Experiment 1, with stress on the frequency-manipulated CVC syllable. It has been suggested that motor programs for stressed and unstressed syllables are independently represented in the syllabary (Crompton, 1981; Levelt, 1989). Thus, in order to keep the basic syllable material between the experiments as consistent as possible, we opted to always have stress on the manipulated syllable.

Design and procedure. The design and procedure were the same as in Experiment 1. We first present and discuss the results of Experiment 2, then those of Experiment 3.

Table 3
Sample Disyllabic Pseudowords in Experiments 2 and 3

Experiment and description	High frequency	Low frequency	High frequency	Low frequency
2: Disyllabic pseudowords with high- and low-frequency first syllables	zɪz.rə	zɪn.rə	ɡɪn.rə	ɡɪz.rə
3: Disyllabic pseudowords with high- and low-frequency second syllables	rə.zɪz	rə.zɪn	rə.ɡɪn	rə.ɡɪz

Results of Experiment 2: Disyllabic Pseudowords With High- and Low-Frequency First Syllables

The raw data were treated the same way as in the previous experiment. Here, 644 (7.9%) trials were treated as errors, and 153 (1.9%) as outliers. The mean voice onset latencies and error rates for Experiment 2 are summarized in Table 4.

The analysis of RTs revealed that disyllabic pseudowords with high-frequency first syllables were produced significantly faster than disyllabic pseudowords with low-frequency first syllables, $t_1(31) = 3.27$, $p = .0026$, $\eta_p^2 = .27$; $t_2(7) = 2.92$, $p = .0223$, $\eta_p^2 = .55$. The analysis of errors yielded no significant effects, $t_1(31) = 1.10$, $p = .2798$, $\eta_p^2 = .04$; $t_2(7) = 1.44$, $p = .1930$, $\eta_p^2 = .23$.

Discussion

Producing disyllabic pseudowords having the frequency manipulation on the first syllables led to significant syllable-frequency effects as well. This result corroborates the results from Experiment 1 and supports the assumption that speakers access precompiled stored syllables during speech planning in English. The finding of an effect for the first syllable is in line with the Dutch (Cholin et al., 2006; Levelt & Wheeldon, 1994) and Spanish (Carreiras & Perea, 2004) findings. But before considering this convergence, we turn to the results of Experiment 3.

Results of Experiment 3: Disyllabic Pseudowords With High- and Low-Frequency Second Syllables

The raw data were treated as in the previous experiments. Here, 535 (6.5%) trials were treated as errors and 162 (2.0%) as outliers. The mean voice onset latencies and error rates for Experiment 3 are summarized in Table 5.

Disyllabic pseudowords were produced significantly faster when the second syllable was of high frequency than when the second syllable was of low frequency, $t_1(31) = 4.02$, $p = .000345$, $\eta_p^2 = .34$; $t_2(7) = 4.21$, $p = .00398$, $\eta_p^2 = .72$. The analysis of errors yielded no significant effects ($t_1 < 1$), $t_2(7) = 1.50$, $p = .1772$, $\eta_p^2 = .24$.

Discussion

The results of Experiment 3 showed significant syllable-frequency effects with high- and low-frequency second syllables. This result again confirms the claim that English speakers use stored syllabic units. Beyond this, though, the results of Experiment 3 contrast with the previous Dutch and Spanish results, which were not associated with second-syllable frequency effects. Finding syllable-frequency effects for first but not for second syllables

Table 5

Mean Voice Onset Latencies (in Milliseconds) and Percentage of Errors for Experiment 3

Frequency	<i>M</i> (<i>SE</i>)	<i>Errors</i> (<i>SE</i>)
High	500 (7)	6.8 (0.4)
Low	514 (8)	6.2 (0.6)
Difference scores	−14 (4)	0.6 (0.6)

implied a syllable-sized planning scope for Dutch and Spanish. The fact that in English both syllables show frequency effects suggests that English's planning scope is larger and may include a stressed syllable and attendant unstressed ones, that is, the planning scope might be *foot*-based. A foot describes a prosodic category or unit (Selkirk, 1982) that includes a sequence of typically two syllables, one of which is metrically stronger than the other one: *trochaic* feet consist of a stressed syllable followed by an unstressed one, whereas *iambic* feet consist of an unstressed syllable followed by a stressed syllable. Dutch and English are languages with a predominant trochaic stress pattern. In these languages it can be argued that the trochaic pattern serves as a default, and hence, the iambic pattern is nondefault, that is, lexicalized (see Booij, 1995; Hammond, 1999; Hayes, 1982; Trommelen & Zonneveld, 1989). What does a default versus nondefault stress pattern have to do with our findings? There may be a difference between the Dutch and English material sets that might be responsible for the different findings for second syllables in these two languages that relates to stress.

The present English study and the former Dutch study used the same item-matching technique (i.e., syllable quartets) as well as the same experimental procedure. The disyllabic pseudowords that were tested for second-syllable frequency effects used CV.CVC targets in English as well as in Dutch and were thus comparable. The divergent findings therefore imply that the differences are in fact due to language. But before we draw this conclusion and discuss details about possible differences between the languages' prosodic units (syllables vs. feet), we need to consider an alternative. There is a systematic difference in the two material sets that were used for the investigation of second-syllable frequency effects: In Dutch, only open CV syllables, such as /ko:byɪ/, served as first syllables, whereas in English, all CV syllables ended in schwa, such as in /rə.gɪn/. Note that the (auditorily presented) pseudowords were always stressed on the manipulated syllable; that is, the disyllabic targets with high- and low-frequency second syllables were presented with an iambic stress pattern (as in /ko:byɪ/ and in /rə.gɪn/) and thus comply with the nondefault stress pattern in Dutch and in English. The fact that all targets were pseudowords that cannot rely on any stored (irregular) stress pattern might have led speakers to fall back on the default (trochaic) stress pattern. If speakers in fact applied this trochaic stress pattern, the Dutch speakers were more likely to find the right base to do so: As already pointed out, the Dutch first syllables consisted of open CV syllables (e.g., /ko:/in/'ko:byɪ/) that receive a long pronunciation and were therefore more likely to attract stress. As such, the speech-planning system might have regarded the stressed syllables as sufficient "chunks" for articulation themselves. This might have been different for the schwa-initial English pseudowords used in Experiment 3. Schwas occur only in unstressed syllables and are often reduced or deleted. Thus, the schwa syllables are unlikely to

Table 4

Mean Voice Onset Latencies (in Milliseconds) and Percentage of Errors for Experiment 2

Frequency	<i>M</i> (<i>SE</i>)	<i>Errors</i> (<i>SE</i>)
High	465 (8)	7.3 (0.5)
Low	475 (8)	8.3 (0.9)
Difference scores	−10 (3)	1 (0.9)

receive stress, and participants might have actually applied the presented, nondefault iambic pattern. Under this scenario, the first syllables might have been even more reduced and were attached to the second syllables, thereby favoring a larger planning scope.

Experiment 4 addresses this interpretation by replicating the third experiment using long/open high-frequency first syllables. If the second-syllable frequency effects for English are due to the use of schwas in the first syllables, the effect should disappear. If they are the result of a generally larger planning scope for English, the second-syllable frequency effect should replicate with long/open first syllables.

Disyllabic Pseudowords With Long/Open High-Frequency First Syllables (Experiment 4)

We created an experimental set that consisted of disyllabic pseudowords with open first CV syllables (diphthongs and long vowels) and stress on the second syllables (see Table 6). Long vowels and diphthongs occupy two slots in the syllable structure (see e.g., Clements & Keyser, 1983), as do vowels in open syllables in Dutch (Booij, 1995). It was tested whether, with these materials, the English participants would opt to start articulation as soon as the first syllable is available or whether they would still postpone the execution of the first syllables until the (frequency-manipulated) second syllable was also available for articulation. If the latter is the case, syllable-frequency effects are expected.

Following the same procedure as in the previous experiments, eight high-frequency CVV syllables were used for the combination with the eight quartets. The CV syllables ranged in frequency from 61.34 to 1,725.08, with an average of 649.85 ($SD = 686.47$) per one million words; the mean number of occurrences in the lexicon per syllable for these items was 250 ($SD = 237$). See Appendix C for a list of the eight syllables and their frequencies (Table C1) as well as for the disyllabic pseudoword material (Table C2). The pairing of two items into one experimental set was the same as in the previous experiment (see Table C2). The numerals to be presented between two experimental items were changed from the previous experiments to 1, 2, 6, and 9 to maintain that there was no onset overlap between numbers and experimental items.

The pseudowords in the current experiment were again presented with stress on second syllables. The spoken syllables were digitized at a sampling rate of 44.1 kHz. They varied in duration from 680 ms to 920 ms, with an average of 832 ms ($SD = 71.31$). There was no difference in duration between high-frequency and low-frequency syllables ($t < 1$).

Method

Participants. Thirty-two native speakers of (American) English took part in Experiment 4. They were taken from the partic-

ipants pool of the Psychology Department at the University of Pennsylvania. They were compensated with \$10 for their participation.

Design and procedure. The design and procedure of Experiment 4 were the same as in the previous experiments.

Results

The raw data were treated in the same way as in the previous experiments. Here, 384 (4.7%) trials were treated as errors and 137 (1.6%) as outliers. The mean voice onset latencies, standard deviations, and error rates for Experiment 4 are summarized in Table 7.

Disyllabic pseudowords were produced significantly faster when the second syllable was high-frequency than when the second syllable was low-frequency, $t_1(31) = 2.06$, $p = .0479$, $\eta_p^2 = .12$; $t_2(7) = 3.29$, $p = .0133$, $\eta_p^2 = .61$. The analysis of errors yielded no significant effects (both $ts < 1$).

Discussion

Disyllabic pseudowords with high-frequency second syllables were again produced significantly faster compared with pseudowords with low-frequency second syllables. Although we hypothesized that the material set for Experiment 3 may not have encouraged a planning scope of syllabic size, the materials for this experiment were “optimized” for a syllabic planning scope. The first syllables contained diphthongs or long vowels (as the first syllables in the Dutch study did) and therefore might have offered better candidates for a syllable-sized planning scope. Nevertheless, English speakers opted to start articulation only after both consecutive syllables were available, as evidenced by the frequency effect for the second syllable. Hence, the contrast between the Dutch and English findings remains.

General Discussion

The article presented four experiments investigating the influence of syllable frequency on the production of mono- and disyllabic pseudowords in English. Syllable-frequency effects were found for all manipulations, providing the first clear support for the notion of stored syllable units that English speakers access during speech production.

Languages with relatively clear boundaries have been shown to be sensitive to syllable-frequency manipulations (French: Laganaro & Alario, 2006; Spanish: Carreiras & Perea, 2004; Dutch: Cholin et al., 2006; Levelt & Wheeldon, 1994, see also Cholin & Levelt, 2009; German: Aichert & Ziegler, 2004, for accuracy data from a study with neurological patients), and it was concluded that these languages predominantly operate on stored syllable-sized

Table 6
Sample Disyllabic English Pseudowords With the Frequency Manipulation on the Second Syllable and Long First Syllables in Experiment 4

High frequency	Low frequency	High frequency	Low frequency
rɔɪ.gɪn	rɔɪ.gɪz	rɔɪ.zɪz	rɔɪ.zɪn
kɑɪ.nɪʃ	kɑɪ.nɪd	kɑɪ.pɪd	kɑɪ.pɪʃ

Table 7
Mean Voice Onset Latencies (in Milliseconds) and Percentage of Errors for Experiment 4

Frequency	<i>M (SE)</i>	<i>Errors (SE)</i>
High	432 (11)	4.8 (0.4)
Low	438 (11)	4.5 (0.4)
Difference scores	-6 (3)	0.3 (0.4)

phonetic units that are accessed during postlexical speech planning. So far, attempts to find evidence for the use of syllabic units in English, in which boundaries are less clear, has been difficult (see e.g., Croot & Rastle, 2004; Cutler et al., 1986), and results from metalinguistic tasks suggest that English speakers may be more sensitive to smaller, subsyllabic units, such as rimes (see e.g., Lee & Goldrick, 2008). Nevertheless, from the current results we can conclude that English, in line with the other tested languages, is also sensitive to syllable-frequency manipulations, suggesting that the underlying phonetic units may be universal, regardless of the syllabic transparency of a language's syllables and its stress properties. But although syllables seem to constitute stored production units regardless of the syllables' transparency in a given language, the transparency might in fact influence the preferred planning scope within different languages (see Table 8 for a systematic comparison of results in English and Dutch).

On the basis of the diverging results of disyllabic pseudowords with the manipulation on the second syllable across languages (see the last three columns of Table 8), we conclude that the comparatively vague syllable boundaries of English lead speakers to prefer a multisyllabic planning scope for multisyllabic pseudowords in some contexts.⁴

But how does one account for the fact that significant frequency effects were found for both first and second syllables in English? After all, if the scope is larger than a single syllable, why should syllable-sized units be stored? For the Dutch results, the case seemed straightforward: The presence of an effect for the first syllable while there was no effect for the second syllable suggests that Dutch speakers use a syllable-sized scope for planning; that is, they start articulation as soon as the first syllable is retrieved and ready to be executed whereas the second syllable is still under construction. The fact that Carreiras and Perea (2004) also reported a null effect for the second syllable but an effect for the first syllable in disyllabic pseudowords suggests that Spanish, a language with transparent syllable boundaries, also plans on the basis of syllables. For English, we offer two possible explanations. First, participants switch between differently sized planning scopes. In Experiment 2 (effect on the first syllable: CVC.CV), the first syllable was rather clear. The participants, who were under pressure to respond quickly, chose a syllable-sized planning scope. In Experiments 3 and 4 (effects on the second syllable: CV.CVC), the syllable boundaries were comparatively less clear. Hence, participants used larger planning scopes, possibly feet. Second, English speakers always use a larger planning scope. This account, though, would question the assumption by Levelt and Wheeldon (1994) that the first syllable is unlikely to evoke frequency effects if speakers initiate articulation only after the final syllable is selected.

Instead, this account assumes some degree of serial processing. First, the first syllable is retrieved in a manner that is sensitive to its frequency, and then the second syllable is retrieved in a like manner.

On the basis of the current results, we cannot choose one explanation over the other, but we would like to propose that, in accordance with both discussed accounts, the relative vagueness of syllable boundaries in English has an impact on the timing of syllabification and syllabary access and, consequently, also on the planning scope. As described in the introduction, syllable access can start as soon as a syllable onset is available, activating all syllables sharing the same onset. In the model of Levelt et al. (1999), verification links checked the activated syllables in the syllabary against the phonological syllables, and, ultimately, the target syllable was selected when the activation threshold was exceeded. A syllable with clear boundaries might pass through these stages relatively faster, or put differently, the sooner a syllable boundary is detected, the easier and therefore faster the syllable located to the left of this boundary can be selected and perhaps independently (from succeeding syllables) be uttered. In this sense, languages with generally clearer syllable boundaries allow for a faster identification of individual syllables.

Following this logic, one might expect generally longer naming latencies for English than for Dutch speakers. However, the material in the respective studies was not matched to allow for such a comparison. Experiment 4 was constructed to be similar to the equivalent Dutch experiment with disyllabic pseudowords that manipulated the frequency of the second syllables. As in the Dutch experiments, the second (high- and low-frequency) syllables carried the stress. As mentioned before, it has been suggested (Crompton, 1981; Levelt, 1989) that there might be two instances for each syllable: a stressed and an unstressed version. Regardless of whether this assumption is correct or reasonable, we cannot rule out that this is not the case. Therefore, we always presented the manipulated syllables with stress and kept this constant throughout all experiments. However, despite presenting the disyllabic pseudowords clearly with stress on the second syllable and a clear instruction to utter these targets as one word, the Dutch participants tended to shift the stress onto the first (open, long) syllables, thereby making them perhaps better stand-alone units that could have served as appropriate articulation units.

As argued earlier, the first syllables with the schwa in Experiment 3 in the present English study might not have offered the right base for treating the initial syllables as stand-alones, but the diphthongs and long vowels that were used as first syllables in Experiment 4 clearly did. The material set of Experiment 4 thus should have allowed English speakers to

⁴ The syllable-frequency effects in Dutch and English are relatively small, that is, they range from 5 to 14 ms. Only the highly controlled material set allowed for the detection of these small effects. Because these studies were not designed to narrow these variable effects down further, we refrain from interpreting the different sizes of the effects across experiments.

Table 8

Summary of the Mean Syllable-Frequency Effects in Mono- and Disyllabic Pseudowords in Dutch and English

Manipulation	Monosyllabic		Disyllabic first		Disyllabic second		
	Dutch	English (Exp 1)	Dutch	English (Exp 2)	Dutch	English (Exp 3)	English (Exp 4)
High	436	464	417	465	435	500	432
Low	445	469	427	475	435	514	438
Difference scores ^a	−9	−5	−10	−10	0	−14	−6

Note. For more information on the Dutch data, please see Cholin et al. (2006). Exp = experiment.

^a Data refer to significance in the analysis over subjects and items.

behave like Dutch speakers. Nevertheless, the frequency effects for second syllables showed that English speakers again encoded both syllables before articulation began. Could one create a material set that would get English speakers to behave like Dutch speakers? All that we can say with certainty is that both of our English experiments, including one whose materials exhibited the key features of the Dutch materials, failed to generate the Dutch pattern. Could one test Dutch and English participants on an even more closely matched material set? Again, we cannot be certain, but it is likely that an even closer matched item set would have to abandon the rigid construction of the syllable quartets and, moreover, to neglect some of the languages' phonology, both of which introduce many uncontrollable factors that would hinder a thorough cross-linguistic comparison. Measuring syllable-frequency effects, in our view, requires the kind of careful control that the quartet method makes possible.

To conclude, all languages that have been tested exhibit syllable-frequency effects. Here, we showed that this is true for English, despite the fact that it has much less transparent syllables than the other languages examined. This suggests that language production generally involves stored syllabic representations. The research further suggests that linguistic differences in stress patterns and syllable transparency may affect (default) planning scopes for articulation. Future research that systematically compares the production of multisyllabic words with syllables of different frequencies may yield further insights into the temporal coordination of planning procedures at the interface between phonological/phonetic encoding and articulation and also further our understanding of the internal makeup of the mental syllabary.

References

- Aichert, I., & Ziegler, W. (2004). Syllable frequency and syllable structure in apraxia of speech. *Brain and Language*, 88, 148–159. doi:10.1016/S0093-934X(03)00296-7
- Alario, F.-X., Costa, A., & Caramazza, A. (2002a). Frequency effects in noun phrase production: Implications for models of lexical access. *Language and Cognitive Processes*, 17, 299–319. doi:10.1080/01690960143000461
- Alario, F.-X., Costa, A., & Caramazza, A. (2002b). Hedging one's bets too much: A reply to Levelt (2002). *Language and Cognitive Processes*, 17, 673–682. doi:10.1080/01690960143000236
- Álvarez, C., Carreiras, M., & Perea, M. (2004). Are syllables phonological units in visual word recognition? *Language and Cognitive Processes*, 19, 427–452. doi:10.1080/769813935
- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical data base* [CD-ROM]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.
- Bachoud-Lévi, A.-C., Dupoux, E., Cohen, L., & Mehler, J. (1998). Where is the length effect? A cross-linguistic study of speech production. *Journal of Memory and Language*, 39, 331–346. doi:10.1006/jmla.1998.2572
- Baus, C., Costa, A., & Carreiras, M. (2008). Neighbourhood density and frequency effects in speech production: A case for interactivity. *Language and Cognitive Processes*, 23, 866–888. doi:10.1080/01690960801962372
- Booij, G. (1995). *The phonology of Dutch*. Oxford, England: Clarendon Press.
- Bradley, D. C., Sánchez-Casas, R. M., & García-Albea, J. E. (1993). The status of the syllable in the perception of Spanish and English. *Language and Cognitive Processes*, 8, 197–233. doi:10.1080/01690969308406954
- Carreiras, M., Álvarez, C., & de Vega, M. (1993). Syllable frequency and visual word recognition in Spanish. *Journal of Memory and Language*, 32, 766–780. doi:10.1006/jmla.1993.1038
- Carreiras, M., & Perea, M. (2002). Masked priming effects with syllabic neighbors in the lexical decision task. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 1228–1242. doi:10.1037/0096-1523.28.5.1228
- Carreiras, M., & Perea, M. (2004). Naming pseudowords in Spanish: Effects of syllable frequency in production. *Brain and Language*, 90, 393–400. doi:10.1016/j.bandl.2003.12.003
- Chen, J.-Y., Chen, T.-M., & Dell, G. S. (2002). Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. *Journal of Memory and Language*, 46, 751–781. doi:10.1006/jmla.2001.2825
- Cholin, J., & Levelt, W. J. M. (2009). Effects of syllable preparation and syllable frequency in speech production: Further evidence for the retrieval of stored syllables at a post-lexical level. *Language and Cognitive Processes*, 24, 662–684. doi:10.1080/01690960802348852
- Cholin, J., Levelt, W. J. M., & Schiller, N. O. (2006). Effects of syllable frequency in speech production. *Cognition*, 99, 205–235. doi:10.1016/j.cognition.2005.01.009
- Clements, G. N., & Keyser, S. J. (1983). *CV phonology*. Cambridge, MA: MIT Press.
- Crompton, A. (1981). Syllables and segments in speech production. *Linguistics*, 19, 663–716. doi:10.1515/ling.1981.19.7-8.663
- Croot, K., & Rastle, K. (2004). Is there a syllabary containing stored articulatory plans for speech production in English? In S. Cassidy, F. Cox, R. Mannell, & S. Palethorpe (Eds.), *Proceedings of the Tenth Australian International Conference on Speech Science and Technology (SST2004)*; pp. 376–381. Sydney, Australia: Macquarie University.
- Cutler, A. (1995). Spoken word recognition and production. In J. L. Miller & P. D. Eimas (Eds.), *Speech, language and communication* (pp. 97–136). San Diego, CA: Academic Press.
- Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's

- differing role in the segmentation of French and English. *Journal of Memory and Language*, 25, 385–400. doi:10.1016/0749-596X(86)90033-1
- Cutler, A., Norris, D., & Williams, J. (1987). A note on the phonological expectations in speech segmentation. *Journal of Memory and Language*, 26, 480–487. doi:10.1016/0749-596X(87)90103-3
- Damian, M. F., & Dumay, N. (2007). Time pressure and phonological advance planning in spoken production. *Journal of Memory and Language*, 57, 195–209. doi:10.1016/j.jml.2006.11.001
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283–321. doi:10.1037/0033-295X.93.3.283
- Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, 27, 124–142. doi:10.1016/0749-596X(88)90070-8
- Dell, G. S., Juliano, C., & Govindjee, A. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science*, 17, 149–195. doi:10.1207/s15516709cog1702_1
- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, 104, 801–838. doi:10.1037/0033-295X.104.4.801
- Dupoux, E. (1993). The time course of prelexical processing: The syllabic hypothesis revisited. In G. Altmann & R. Shillcock (Eds.), *Cognitive models of speech processing* (pp. 81–114). Hillsdale, NJ: Erlbaum.
- Gordon, J. K., & Dell, G. S. (2001). Phonological neighborhood effects: Evidence from aphasia and connectionist models. *Brain and Language*, 79, 21–23.
- Hammond, M. (1999). *The phonology of English*. Oxford, England: Oxford University Press.
- Hayes, B. (1982). Extrametricality and English stress. *Linguistic Inquiry*, 13, 227–276.
- Hendriks, H., & McQueen, J. (Eds.). (1996). *Annual report 1995*. Nijmegen, the Netherlands: Max Planck Institute for Psycholinguistics.
- Jeschienak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 824–843. doi:10.1037/0278-7393.20.4.824
- Kahn, D. (1980). Syllable-structure specifications in phonological rules. In M. Aronoff and M.-L. Kean (Eds.), *Juncture* (pp. 91–105). Saratoga, CA: Anna Libri.
- Laganaro, M., & Alario, F.-X. (2006). On the locus of the syllable frequency effect in language production. *Journal of Memory and Language*, 55, 178–196. doi:10.1016/j.jml.2006.05.001
- Lee, Y., & Goldrick, M. (2008). The emergence of sub-syllabic representations. *Journal of Memory and Language*, 59, 155–168. doi:10.1016/j.jml.2008.03.002
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. M. (2002). Picture naming and word frequency. *Language and Cognitive Processes*, 17, 663–671. doi:10.1080/01690960143000443
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–75. doi:10.1017/S0140525X99001776
- Levelt, W. J. M., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition*, 50, 239–269. doi:10.1016/0010-0277(94)90030-2
- MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences*, 21, 499–546. doi:10.1017/S0140525X98001265
- Mehler, J., Dommergues, J., & Frauenfelder, U. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, 20, 298–305. doi:10.1016/S0022-5371(81)90450-3
- Meyer, A. S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language*, 29, 524–545. doi:10.1016/0749-596X(90)90050-A
- Meyer, A. S. (1991). The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language*, 30, 69–89. doi:10.1016/0749-596X(91)90011-8
- Meyer, A. S., Belke, E., Häcker, C., & Mortensen, L. (2007). Use of word length information in utterance planning. *Journal of Memory and Language*, 57, 210–231. doi:10.1016/j.jml.2006.10.005
- Meyer, A. S., Roelofs, A., & Levelt, W. J. M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language*, 48, 131–147. doi:10.1016/S0749-596X(02)00509-0
- Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, 17, 273–281. doi:10.1080/17470216508416445
- Pulgram, E. (1970). *Syllable, word, nexus, cursus*. The Hague, the Netherlands: Mouton.
- Roelofs, A. (1997a). Syllabification in speech production: Evaluation of WEAVER. *Language and Cognitive Processes*, 12, 657–693. doi:10.1080/016909697386655
- Roelofs, A. (1997b). The WEAVER model of word-form encoding in speech production. *Cognition*, 64, 249–284. doi:10.1016/S0010-0277(97)00027-9
- Schiller, N. O., & Costa, A. (2006). Activation of segments, not syllables, during phonological encoding in speech production. *Mental Lexicon*, 1, 231–250. doi:10.1075/ml.1.2.04sch
- Schiller, N. O., Fikkert, P., & Levelt, C. C. (2004). Stress priming in picture naming: An SOA study. *Brain and Language*, 90, 231–240. doi:10.1016/S0093-934X(03)00436-X
- Schiller, N. O., Meyer, A. S., Baayen, R. H., & Levelt, W. J. M. (1996). A comparison of lexeme and speech syllables in Dutch. *Journal of Quantitative Linguistics*, 3, 8–28. doi:10.1080/09296179608590060
- Schriefers, H., & Teruel, E. (1999). Phonological facilitation in the production of two-word utterances. *European Journal of Cognitive Psychology*, 11, 17–50. doi:10.1080/713752301
- Selkirk, E. O. (1982). The syllable. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representations: Part II* (pp. 337–383). Dordrecht, the Netherlands: Foris.
- Spencer, A. (1996). *Phonology*. Oxford, England: Blackwell.
- Taft, M., Álvarez, C. J., & Carreiras, M. (2007). Cross-language differences in the use of internal orthographic structure when reading polysyllabic words. *Mental Lexicon*, 2, 49–63. doi:10.1075/ml.2.1.04taf
- Treiman, R. (1983). The structure of spoken syllables: Evidence from novel word games. *Cognition*, 15, 49–74. doi:10.1016/0010-0277(83)90033-1
- Treiman, R., & Danis, C. (1988). Syllabification of intervocalic consonants. *Journal of Memory and Language*, 27, 87–104. doi:10.1016/0749-596X(88)90050-2
- Treiman, R., & Zukowski, A. (1990). Toward an understanding of English syllabification. *Journal of Memory and Language*, 29, 66–85.
- Trommelen, M., & Zonneveld, W. (1989). *Klemtoon en metrische fonologie* [Stress and metrical phonology]. Muiderberg, the Netherlands: Coutinho.
- Vitevitch, M. S. (1997). The neighborhood characteristics of malapropisms. *Language and Speech*, 40, 211–228.
- Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 735–747. doi:10.1037/0278-7393.28.4.735
- Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and Language*, 68, 306–311. doi:10.1006/brln.1999.2116

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

Wheeldon, L. R., & Levelt, W. J. M. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language*, 34, 311–334. doi:10.1006/jmla.1995.1014

Appendix A

Experiment 1

Table A1
Syllable Quartets

Quartet no.	High-frequency syllables			Low-frequency syllables		
	Syllable	Summed frequency	No. of occurrences	Syllable	Summed frequency	No. of occurrences
1	gm	133.18	7	gɪz	1.9	16
1	zɪz	619.89	645	zɪn	2.01	3
2	fæk	318.47	60	fæl	1.28	7
2	kæl	73.30	53	kæk	2.57	5
3	nɪʃ	164.24	56	nɪd	7.71	15
3	pɪd	12.57	26	pɪʃ	6.54	27
4	fɪd	66.87	2	fɪk	4.36	1
4	fɪk	334.97	107	fɪd	4.97	6
5	vəl	64.75	31	vɑŋ	0	1
5	kɑŋ	100.06	58	kəl	4.97	7
6	lɪf	61.45	20	lɪg	0.39	10
6	sɪg	227.37	58	sɪf	2.57	4
7	jʌd	93.35	3	jʌs	5.59	3
7	fʌs	437.32	225	fʌd	8.10	5
8	jes	774.24	4	jen	1.34	1
8	ben	68.32	29	bes	0	1

Table A2
Pairing of Syllables Into Frequency-Homogeneous Sets in Experiment 1

Set no.	High-frequency sets	Low-frequency sets
1	gm-ben	gɪz-bes
2	zɪz-jes	zɪn-jen
3	fæk-nɪʃ	fæl-nɪd
4	kæl-pɪd	kæk-pɪʃ
5	fɪk-kɑŋ	fɪd-kəl
6	fɪd-vəl	fɪk-vɑŋ
7	jʌd-sɪg	jʌs-sɪf
8	fʌs-lɪf	fʌd-lɪg

(Appendices continue)

Appendix B

Experiment 2 and 3

Table B1

CV Syllables That Served as First (Experiment 2) and as Second (Experiment 3) Syllables in the Disyllabic Pseudowords

Syllable	Summed frequency	No. of occurrences
bə	5,841	841
də	2,111	882
gə	471	469
lə	2,090	1,117
nə	2,302	963
mə	2,327	1,142
rə	7,446	1,821
tə	25,401	2,700

Note. CV = consonant–vowel.

Table B2

Materials for Experiments 2 and 3

Quartet no.	Experiment 2		Experiment 3	
	High-frequency sets	Low-frequency sets	High-frequency sets	Low-frequency sets
1	gɪnrə	gɪzrə	rəɡɪn	rəɡɪz
1	zɪzrə	zɪnrə	rəzɪz	rəzɪn
2	fækbbə	fælbə	bəfæk	bəfæl
2	kælbə	kækbə	bəkæl	bəkæk
3	nɪʃgə	nɪdgə	ɡəniʃ	ɡəniɪd
3	pɪdgə	pɪʃgə	ɡəpiɪd	ɡəpiʃ
4	ʃɪdlə	ʃɪklə	ləʃɪd	ləʃɪk
4	fɪdlə	fɪklə	ləfɪk	ləfɪd
5	vəltə	vəŋtə	təvəl	təvəŋ
5	kəŋtə	kəltə	təkəŋ	təkəl
6	lɪfnə	lɪgnə	nəlɪf	nəlɪɡ
6	sɪgnə	sɪfnə	nəsɪɡ	nəsɪf
7	jədmə	jəsmə	məjəd	məjəs
7	ʃəsmə	ʃədmə	məʃəs	məʃəd
8	jesdə	jendə	dəjes	dəjen
8	bendə	besdə	dəben	dəbes

(Appendices continue)

Appendix C

Experiment 4

Table C1

CV Syllables That Served as First Syllables in the Disyllabic Pseudowords in Experiment 4

Syllable	Summed frequency	No. of occurrences
rɔɪ	101.28	12
naɪ	393.58	160
daʊ	61.34	56
lɔɪ	61.40	23
pəʊ	493.41	221
mɑː	684.47	351
teɪ	1,725.08	592
keɪ	1,678.27	583

Note. CV = consonant–vowel.

Table C2

Materials for Experiment 4

Quartet no.	High-frequency sets	Low-frequency sets
1	rɔɪgɪn	rɔɪgɪz
1	rɔɪzɪz	rɔɪzɪn
2	pəʊfæk	pəʊfæl
2	pəʊkæl	pəʊkæk
3	keɪnɪf	keɪnɪd
3	keɪpɪd	keɪpɪf
4	lɔɪfɪd	lɔɪfɪk
4	lɔɪfɪk	lɔɪfɪd
5	teɪvəl	teɪvən
5	teɪkəʊ	teɪkəl
6	næɪlɪf	næɪlɪg
6	næɪsɪg	næɪsɪf
7	mɑːjəd	mɑːjəs
7	mɑːʃəs	mɑːʃəd
8	dəʊjes	dəʊjen
8	dəʊben	dəʊbes

Received June 24, 2009

Revision received June 26, 2010

Accepted June 29, 2010 ■