Proximate Units in Word Production: Phonological Encoding Begins with Syllables in Mandarin Chinese but with Segments in English

Padraig G. O'Seaghdha,
Lehigh University, USA

Jenn-Yeu Chen, and
National Cheng Kung University, Taiwan

Train-Min Chen
National Cheng Kung University, Taiwan

Abstract

In Mandarin Chinese, speakers benefit from fore-knowledge of what the first syllable but not of what the first phonemic segment of a disyllabic word will be (Chen, J.-Y., Chen, T.-M. & Dell, 2002), contrasting with findings in English, Dutch, and other Indo-European languages, and challenging the generality of current theories of word production. In this article, we extend the evidence for the language difference by showing that failure to prepare onsets in Mandarin (Experiment 1) applies even to simple monosyllables (Experiments 2-4), and confirm the contrast with English for comparable materials (Experiments 5, 6). We also provide new evidence that Mandarin speakers do reliably prepare tonally unspecified phonological syllables (Experiment 7). To account for these patterns, we propose a language general *proximate units principle* whereby intentional preparation for speech as well as phonological-lexical coordination are grounded at the first phonological level below the word at which explicit unit selection occurs. The language difference arises because syllables are proximate units in Mandarin Chinese, whereas segments are proximate in English and other Indo-European languages. The proximate units perspective reconciles the aspiration toward a language general account of word production with the reality of substantial cross-linguistic differences.

Keywords

phonological encoding; proximate units; cross-linguistic analysis; syllable production; speech planning

Contact information: Padraig G. O'Seaghdha Department of Psychology Lehigh University 17 Memorial Drive East Bethlehem, PA 18015-3068 pat.oseaghdha@lehigh.edu Phone: 610 758-4526 Fax: 610 758-6277 Jenn-Yeu Chen Institute of Cognitive Science National Cheng Kung University 1 University Drive Tainan, Taiwan 701 psyjyc@mail.ncku.edu.tw.
The first and second authors contributed equally to this study.

Proximate Units in Word Production: Phonological Encoding Begins with Syllables in Mandarin Chinese but with Segments in English

It has become virtually doctrinal in language production theory that phonemic segments provide the common currency of phonological encoding (e.g., Bock 1991; Dell 1986, 1995; Levelt 1989; Levelt et al., 1999; Rapp and Goldrick, 2000). Though they disagree in many details, these theories concur that phonological segments are retrieved from the lexicon and linearized in a syllabified organization that guides articulation. Yet it is also widely acknowledged that this consensual view may not apply without qualification to all languages, in particular to Chinese languages in which entire syllables are available at lexical retrieval (e.g., Chen et al., 2002; Cholin, Levelt & Schiller, 2004; Levelt, 2001). Thus, the field adheres to theories that emerged in the context of European languages despite the acknowledged challenges to their generality. In this article, we show that differences between Chinese and Indo-European languages are more deep-seated than has been previously recognized, and we argue that these differences require more than minor adjustments to existing theories. We outline a response to this challenge that focuses on the starting point of phonological retrieval. We propose that the first selectable phonological units below the level of the word, which we call *proximate units,* vary across languages and are often pivotal in situations such as advance planning and partial preparation that involve continued coordination of phonological ingredients with their lexical origins. Proximate units also constrain the phonological codes that are assembled to guide the articulation of stretches of speech. Thus our approach highlights the critical juncture between words and phonology, but also addresses the sequelae of proximate unit retrieval.

We first describe the motivation of our idea in the literature on production of Chinese words. Then we present our proximate units proposal in more detail. Our hypothesis that syllables are proximate units in Mandarin, but not in English leads to a two fronted empirical agenda. On one hand we provide a more stringent test of the language difference using simple monosyllables in Mandarin and English. On the other we consolidate the evidence that the first phonological syllables of Mandarin disyllables are indeed functional planning units.

Phonological Encoding in Mandarin Chinese

Descriptive and linguistic studies of Chinese phonology suggest that word form construction depends systematically on syllable units (Cheng, 1973; Kuo, 1994). In Mandarin, syllables are simple units, most of them CV, CVV or CVC. There are only two codas (/n/, /ng/) and resyllabification is nonexistent. There are four primary tones and each syllable requires a tone value. Except that tone is sometimes contextually adjusted (see e.g., M. Y. Chen, 2000), what you retrieve from the Mandarin lexicon is what you say. Mandarin words are constructed from a rather small inventory of syllable units, each syllable being homophonic with multiple free or bound morphemes (Packard, 2000). There are about 400 syllables, disregarding tone, but about 1200 tonal syllables. Throughout this article we use the term *tonal syllable* to signify a syllable with a specified tone and *atonal syllable* for a syllable with indeterminate tone.

Studies of naturally occurring speech errors, computational demonstrations, and experimental evidence all support the inference that Mandarin Chinese uses syllable units as fundamental building blocks of word form encoding. The speech error evidence is particularly crucial because it has been established that units that figure prominently in speech errors are the pivotal ones in phonological encoding. Bock (1991) and Dell (1995) assessed the most noticeable and perhaps least ambiguous errors, exchanges, and found that the vast preponderance of English phonological errors belong to two categories, words or morphemes on the one hand, and segments on the other. In contrast, a reanalysis of Chen's (1993) small corpus of Mandarin Chinese errors shows three categories, word, syllable, and segment. Of 211 errors, there were

substantial numbers of both syllabic (35) and subsyllabic (77) errors. Considering only exchanges, 16 of 41 were syllabic and only 4 were segmental, with the rest (21) involving whole words. This is consistent with the view that words and syllables are directly involved in the linearization of the speech plan, and that segments are at one remove from this process. Because the Chen (1993) sample is quite small the error counts must be viewed only as suggestive estimates. Nonetheless, it is clear a) that in contrast to English, syllable exchanges are well represented in Mandarin, and b) that syllable exchanges are more frequent than segment exchanges in Mandarin.

With a separate and larger error collection, Chen (2000) showed that both tonal and atonal syllables are error units in Mandarin Chinese. The fact that atonal syllables slip [e.g., wu4-*cha*1 ('error') jian3-*xiao*3 ('reduce') → wu4-xiao1 (a nonword anticipating the second syllable of the next word but preserving the tone of the replaced syllable)], suggests strongly that tone is represented separately from the segmental ingredients (Chen, 1999), and that syllables are selected phonological units. Contrast to this the paradoxical situation in English where syllables constrain segmental errors but very rarely appear as error units (Shattuck-Hufnagel, 1983; Dell, 1986; see Sevald, Dell, & Cole, 1995). These findings already suggest that word form encoding in these two languages is divergent.

Computational evidence also supports the language contrast. Chen, Dell, and Chen (2004) trained a recurrent sequential network to predict the next phonological segment of Mandarin and English translations of the same texts presented in an undemarcated stream. They showed that vivid syllable structure emerged from the sound sequence of Mandarin Chinese. In contrast, word boundaries were much more salient than within-word syllable boundaries in English. That is, given only a continuous stream of segments, the network discovered the much greater syllabic order in Mandarin.

Turning to experimental research, form preparation studies that we review in more detail below, suggest that Mandarin speakers benefit when a set of words produced in random order shares initial syllables, whereas speakers of European languages but *not* Mandarin speakers benefit from knowing the initial sounds (hereafter *onsets*) of sets of words (Chen, Chen & Dell, 2002; Meyer, 1990, 1991). The null effect of shared onsets in Mandarin suggests that form preparation reflects *intentional orientation* to phonological components and not mere exposure to or use of them.

Consistent with both the error data and the form preparation findings, in a word pair recitation task that induces strong similarity-based phonological competition, disyllable pairs sharing first syllables experienced strong competition in Mandarin (Chen, O'Séaghdha & Liu, 2007). Competition was present when tone varied though it was greater when it was the same. Together with the other evidence already cited, we suggest that competition in Mandarin involves syllable units that are selected, and sometimes miss-placed, wholistically.

Word onsets not only show benefits in European languages when they are shared in a set of words to be produced (Meyer, 1990, 1991), they also produce benefits when presented via distractors during picture naming (e.g., Meyer & Schriefers, 1991; Jescheniak & Schriefers, 2001). However, using another Chinese language, Cantonese, Wong and H. C. Chen (2008) recently reported a surprising lack of effect for visual or auditory distractors sharing onsets. In contrast, there were clear benefits of atonal syllable distractors, suggesting that the syllable and not the initial segment was available early in phonological encoding.

Similarly, atonal syllables but not their onsets are effective masked primes in Mandarin (Chen, Lin & Ferrand, 2003; Chen, J.-Y., Chen, T.-M. & O'Seaghdha, 2009) whereas the case for distinctly syllabic priming in European languages is not proven (Schiller, 1998; Schiller, Costa & Colomé, 2002; Perret, Bonin & Meot, 2006; but see Ferrand, Segui & Grainger, 1996).

Because the primes are by design activated but not selected for production in masked priming, this suggests that in Mandarin syllable units are activated even when they are not under consideration for production.

Taken together, these cross-language differences point to a fundamental difference in how syllable and segment units are engaged in Chinese and European languages. Specifically, we hypothesize that Mandarin speakers retrieve entire syllable units and use them to guide further phonological encoding, whereas speakers of European languages retrieve the segments of words or morphemes without explicit syllable mediation. We now turn to look in some detail at the form preparation task evidence for this contrastive configuration of Chinese-language and European-language word encoding procedures.

Form Preparation Across Languages

In the form preparation task, also known as implicit priming in production (Meyer, 1990, 1991), participants study small sets of three or four prompts and response words (see Appendix tables for examples from our own Mandarin and English experiments). Then, they name the response words as quickly as possible as the prompts unpredictably appear. The crucial aspect of the design is that the context in which the items in a set appear is either *homogeneous*, meaning that the items share some ingredients, or *heterogeneous,* meaning that they have nothing systematically in common. If participants name words more quickly in the homogeneous context, that is de facto evidence that the shared ingredient has contributed to advance preparation. Thus the form preparation task can be used to discover what units are engaged in advance planning of word production.

Meyer reported several key findings for Dutch. There is a benefit of knowing the onset of a set of words, whether the words are monosyllabic (Meyer, 1991) or disyllabic (Meyer, 1990). Knowing that words share later components when onsets differ is not beneficial (Meyer, 1991), presumably because preparation must engage the sequential assembly of the presented word. Knowing more consecutive beginning segments of a word increases effect size. Also, advance knowledge of metrical properties such as number of syllables or primary stress location is not beneficial by itself, but variability in these properties may undermine the segmental effects (Roelofs & Meyer, 1998). In general, we may conclude that individual segments are primary planning units, and that they are deployed sequentially in the context of metrical properties (Levelt et al., 1999).

In contrast, Chen et al. found that Mandarin speakers benefit from knowing the first syllable of a disyllable, but not just the onset of the first syllable (Chen et al., 2002). The syllable benefit accrued even if the first syllables varied in tone (atonal syllables), but was greater if tone was shared (tonal syllables). It did not matter whether the tonal syllables pertained to a unique morpheme, suggesting that the functional unit is phonological not morphological (see also Chen, T.-M. & Chen, J.-Y., 2006; Chen, J.-Y. & Chen, T.-M., 2007). Moreover, just as metrical properties alone did not help Dutch speakers (Roelofs & Meyer, 1998), and consistent with Chen's speech error analyses, tone alone provided no benefit, suggesting that tone is not a free-standing planning unit (Chen et al., 2002).

Taken together, these findings point to commonalities and differences in the processes underlying word form encoding in European and Chinese languages. Current English- and Dutch-centric theories differ in specifics but agree that word form encoding is controlled at the level of words or morphemes, and that segments are disbursed in parallel and assembled sequentially in the syllabified organization that guides production (Levelt, Roelofs & Meyer, 1999; Sevald & Dell, 1994; see O'Seaghdha & Marin, 2000). In our view, Mandarin is fundamentally different in that words are built from syllable units, and so segments are not directly bound to words. Rather, syllables actuate phonological encoding at the level of

segments. In short, consistent with the speech error evidence, atonal syllables are explicitly selected as phonological chunks in Mandarin and then specified segmentally and tonally. This means that the retrieved proximate units and the specified prearticulatory units correspond much more closely than they do in European languages, a point we return to in the General Discussion.

Although the focus of our article is on Chinese and European languages, our proposal is general, and so it is not centered on a binary choice between syllables and segments as proximate units. Fortunately, Japanese provides another relevant case in which proximate units appear to be neither syllables nor segments.

Japanese Morae—Unlike Indo-European and Chinese languages, Japanese is moraic, meaning that subsyllabic moraic units are distinct production units (see Kubozono, 1989). Japanese morae vary in composition, some being single segments but most comprising a consonant and vowel (CV). Given their evident role in speech perception (e.g., Otake et al, 1993) and production, we suggest that morae are the likely proximate units in Japanese. If so, onsets of word initial CV morae would be inaccessible to form preparation just as the onsets of Mandarin words are.

In a set of form-preparation experiments, Kureta, Fushimi and Tatsumi (2006) examined tri-moraic words beginning with a consonant-vowel (CV) mora and continuing with a second mora that was either a vowel or a nasal consonant. In the present context, the key question is whether the onsets of the initial CV morae yielded form preparation benefits. Consistent with the Chen et al. findings for syllable onsets, they did not (Kureta et al., Experiment 1). However a clear benefit was observed for the word initial CV mora unit, and additional benefits accrued when the second mora, which sometimes comprised a single consonant, could be conjoined with the first (Kureta et al., Experiment 3). Thus, whether the functional unit is a mora in Japanese or an entire syllable in Mandarin, it seems to preempt preparation of subordinate segmental components, specifically onset consonants. This evidence from a quite different phonological system corroborates our proposal that onset preparation may indeed be precluded when multi-segmental units are the first resort of word form encoding.

Proximate Units

Proximate units are *the first selectable phonological units below the level of the word/ morpheme*. These units vary across languages, but according to our *proximate units principle* (O'Seaghdha & Chen, 2009) they serve several crucial roles in word encoding regardless of the type of unit. The hallmark of proximate units is that they are selected in the initial stage of phonological encoding. As reviewed earlier, they are therefore prone to selection errors, including exchange errors. For Mandarin, we propose that proximate units comprise syllable chunks rather than phonemic segments, and so addressing onsets is not reflexive. For English, proximate units comprise phonological segments that are indexed left to right, thus making onsets immediately accessible to speakers.

O'Séaghdha & Chen (2009) discussed additional roles of proximate units, focusing on situations of partial preparation or advance planning where what is initially retrieved from the lexicon is likely to be in play. We proposed that proximate units channel phonological activation in TOT states and so constrain subjective reports of partial phonology. They also constrain what a speaker will spontaneously conceive as a "word beginning", and thus what speakers of different languages address intentionally in response to that query. In the present context, Mandarin speakers naturally intend to produce syllables, perhaps to the exclusion of subsyllabic ingredients. In contrast, English speakers intentionally address word onsets, or combinations of linked word-initial segments such as consonant-vowel, consonant-vowel-consonant, and so on.

By focusing on the units that are selected first during encoding we are setting aside consideration of other manifestations of syllabicity in European languages (see Laganaro & Alario, 2006, for review). For example, there are proposals that syllables are represented only or primarily structurally in European languages (see Costa & Sebastian, 1998; Ferrand & Segui, 1998; Sevald, Dell & Cole, 1995), and that syllables are retrieved at a late post-phonological, prearticulatory stage (Levelt et al., 1999; Cholin, Levelt & Schiller, 2004; Cholin, 2008; see also Laganaro & Alario, 2006). Our assertion that syllables are not initially retrieved as chunks in English and related languages does not speak to these and other possibilities and so should not be read as asserting that syllabicity plays no role in phonological encoding in European languages.

The form preparation task can be used to test our ideas concerning the initial encoding of Mandarin and English. If our proposals are correct, a Mandarin speaker given a set of words that share the first segment will view these as involving several distinct syllables. In contrast, an English speaker with such a set of words in mind is immediately able to isolate the onset and thus expedite completion of word preparation when the specific word to be produced is provided. As noted earlier, an English speaker may also concatenate several contiguous segments of a word beginning, but apparently without regard to their mapping to the word's citation syllabification (Meyer, 1990; Schiller, 1998). These findings are consistent with our proposal that the *proximate* phonological units, those that are initially selected, are different in the two languages.

Testing Onset Accessibility in Mandarin

The proximate units proposal predicts that word onsets are generally accessible in languages such as English or Dutch under circumstances where they are not in Chinese languages. Experiment 1 replicated the disyllable onset test of Chen et al. (2002, Experiment 5) with simplified materials. This sets the stage for the following monosyllable experiments in which the same first syllables and design as in Experiment 1 were deployed.

Assuming that syllables are in fact pivotal in Mandarin phonological encoding, an important question is whether the syllabic organization is intrinsic or whether it may be instead a product of the sequencing requirements of multisyllabic words. Under the latter contingency, speakers might assemble syllables similarly in English and Mandarin, but because Mandarin syllables are so regular and systematic, speakers could use them strategically for planning purposes. To exclude this possibility, we used words of only one syllable in most of the experiments (Experiments 2-6) in this article. If syllables are proximate units and not merely units of convenience, encoding a Mandarin monosyllable is different than encoding an English one in that Mandarin requires selection of the phonological syllable prior to retrieval of its segments, whereas English directly retrieves segments.

Figure 1a outlines the postulated sequential steps of encoding a Mandarin monosyllable in a standard notation that assumes separation of content and structure. First a lexicalized concept, such as "virtue", is activated. Activation flows to the corresponding abstract word node or *lemma,* which in turn links to phonological content and structure. (The notation de2 at the word level means that the syllable */də/* and its tone value are indexed to the word, not that they are realized at this level). Phonological content comes bundled in syllables which need to be sequenced in complex words or merely enacted in the case of monosyllables. When a syllable is to be produced, it is selected by linking it to the corresponding unit in the structure network and a syllable frame is retrieved. Then the segments of the syllable are disbursed in parallel and linked sequentially to positions in the syllable frame. Metrical tone is also specified at this point.

If speakers do not normally access syllable-internal components until the later stage of syllable encoding, then the fact that the monosyllables in a form preparation set share an onset may provide no benefit. Conversely, if the focus on syllables is strategic, when only monosyllables are to be produced, the focus of preparation may be shifted downward to segments, making production of monosyllables functionally indistinguishable from that in Indo-European languages.

The parallel English monosyllable example (Figure 1b) does not have an explicit syllable level, so there is no distinction between word/morpheme and syllable for monosyllables. When a monosyllabic word such as "day" is to be produced, its segments are activated, selected, linked directly to positions in the word-shape frame, and default stress (not shown in the Figure) is applied. The important point is that, in contrast to Mandarin, syllables are not selectable chunks and so they do not appear at all in the monosyllable case.

We examined these alternatives by testing Mandarin speakers with simple monosyllables and by comparing the performance of English speakers with equivalently simple monosyllabic words. Experiments 2-4 tested for onset preparation of Mandarin monosyllables in conditions that systematically removed possible obstacles to the expression of the effect. In Experiment 2, the prompt monosyllabic word (e.g., 觸 /chu4/ 'touch') and the response (摸 /mo1/ also 'touch') comprised a disyllabic compound (觸摸 /chu4-mo1/ 'to touch'). In Experiment 3, they did not. In Experiment 4, to remove any possibility of lexicalizing the prompt-response combinations, the prompts were nonlexical symbols (e.g., @). To anticipate our findings, there was no onset preparation effect in any of these Mandarin experiments.

Our null results with Mandarin onsets contrast with positive effects in equivalent English experiments and the benefit of preparing atonal first syllables in Mandarin. We report two English language experiments with comparably simple monosyllable materials in which we confirm the benefit of onset preparation. Finally, Experiment 7 provides a rigorous test of both the presence of an atonal syllable benefit and the absence of an onset preparation benefit with exactly the same words counterbalanced over conditions and participant groups.

Following standard procedure (Meyer, 1990; Chen et al., 2002), all experiments included three repeating blocks of each set. This design feature provides data stability, while also enabling assessment of adaptations during testing. For example, it may take time for speakers to become familiar enough with the experimental context to respond optimally in homogeneous conditions. Mandarin speakers who do not routinely attend to onsets could become attuned to them in experiments where all responses are monosyllabic. Likewise, the procedures of syllable preparation may take time to stabilize. In either case, form preparation effects would not appear at the beginning of the experiment but rather emerge over blocks.

General Method: Mandarin Experiments 1-4

The method in these experiments was the same except as specified later for each experiment, so we present the shared method here. In each case, participants received four homogeneous and four heterogeneous sets, and each set contained four items.

Participants

Each experiment used 12 undergraduate students from National Cheng Kung University. They were all native Mandarin Chinese speakers and were paid for participation.

Apparatus

The experiments were run on a PC controlled by a customized C program. Responses were registered with a voice key interfaced to the computer. A Cowon iAudio 6 multi-media displayer was used to record the participants' responses for later coding and scoring.

Materials

The core materials were four sets of four monosyllables, each set having a different onset, /m/, /d/, /sh/, *or* /l/. These core materials were deployed in each of Experiments 1-4 (see Appendix). To maximize the potential for preparation of onsets, they were adapted from the materials in Chen et al. (2002) to be as simple and unambiguous as possible. Any complex syllables (e.g., containing diphthongs or glides), were reduced to simple consonant-vowel (CV) structures. If the character corresponding to a target had more than one potential pronunciation, for example with alternative tones, it was replaced with an unambiguous character. The four monosyllables in each set were the targets for production in Experiments 2-4 and were the first syllables of disyllabic words in Experiment 1. Each target was paired with a prompt. In Experiment 1, the prompts and targets were associated disyllabic words. In Experiment 2, the prompts were monosyllables that formed forward (7 items), backward (7 items), or bidirectional (2 items) disyllabic compounds with the target syllables (e.g., from the /m/ set, chu4-mo2 觸--摸 means 'touch', ma2-que4 麻--雀 means 'sparrow', and feng1-mi4 蜂--蜜 means 'honey' when read forward and 'bee' when read backward). The backward items were reversed during presentation. The prompts in Experiment 3 were associative monosyllables that did not form compounds with the targets. That is, both prompts and targets were monosyllabic words. In Experiment 4, the prompts were nonlexical symbols.

As shown in Appendix Table 1 for Experiment 1, there were 4 fully counterbalanced sets of 4 heterogeneous and 4 homogeneous targets in each experiment. In the homogeneous condition, the first syllables of the response words within a set all shared the same onset consonant but differed in their vowels and tones. For example, '摸' [**m**o1], '麻' [**m**a2], '牡' [**m**u3], '蜜' [**m**i4], have the same onset /m/ but have different vowels and tones. In the heterogeneous condition, the materials were rearranged so that no segments were shared among members of a set, for example, the first syllables were '摸' [mo1], '賭' [du3], '實' [shi2], '律' [lv4].

Design

Experiments 1-4 all used the same experimental design, adopted from Chen et al., (2002). The primary manipulation, Context, was whether the sets were homogeneous or heterogeneous. Because the context experienced first could influence responding, the Sequence of contexts was counterbalanced across participants. Half of the participants received homogeneous contexts first and half received heterogeneous contexts first. Each of the four Sets in a context was presented in random order before switching to the other context. In a presentation block, the prompt-response items in a set were presented in random order four times (Repetitions) and the entire Block procedure was replicated three times before moving to the next set. Altogether, each participant received 384 trials (3 blocks* 2 contexts* 4 sets* 4 items* 4 repetitions). Except for Sequence of contexts, which was counterbalanced, all factors were within-subjects.

Procedure

We used the standard associative procedure introduced by Meyer (1990). This involved two phases, a learning phase that preceded each set, and a testing phase. During the learning phase, the program cued the experimenter as to what set was scheduled next, and participants were then shown an index card including the four prompt-response pairs to be tested. They learned

(or re-learned) the associations between the prompts and responses in the set until they indicated that they were ready for the production task. Then they proceeded to the testing phase.

During the testing phase, each trial began with a 200 ms, 1000-HZ warning tone and two short dashed lines flanking the blank space at the center of the screen where the prompt would appear. The prompt appeared in the previously flanked space 600 ms after the offset of the tone and stayed on the screen for 150 ms. Participants were instructed to say the target word aloud, as quickly and accurately as possible, in response to the prompt. If there was no response within 1000 ms, a timeout was signaled by a 200 ms 500-HZ tone. The next trial began 200 ms after the response or the timeout signal.

One heterogeneous practice set with items not used in the experimental sets was given before the experiment began. The participants were seated about 60 cm from the screen. Each character measured 1.6 cm in height and 1.1 cm in width. It took about 45 minutes to complete a session.

Experiment 1: Disyllable Onsets

In Experiment 1, we replicated Experiment 5 of Chen et al. (2002) with the slightly simplified materials described above. The experiment provides another test of the shared onset conditions with optimized materials, and also provides the baseline of comparison of Experiments 2-4 in which the first syllables of the target words were tested. The core materials described above comprised the first syllables of the disyllabic response words as shown in Appendix Table 1.

Results

Errors including incorrect responses, failures to respond, stutters, and filled pauses preceding responses were segregated from the valid responses. Table 1 shows the mean of the valid response times and error rates (in parentheses) for the four sets of targets in each homogeneous and heterogeneous set as well as the overall effect, 3 ms. The error rate was 3% both in the homogeneous and in the heterogeneous contexts.

Analyses of variance were conducted by subjects and by items on correct response times. In the by-subjects analyses, responses to the same item within a block were averaged. Heterogeneous items were regrouped by Set. We report analyses by subjects and by items separately.

Overall, the 3 ms onset preparation effect was null: $F1$ (1, 10) < 1, MSE = 1540. The preparation effect did vary significantly across sets, ranging between 12 ms and −10 ms [Context × Set: $F1$ (3, 30) = 3.30, MSE = 577, $p < .05$], but none of the individual set effects was statistically reliable. We interpret this variation as largely random.

Based on previous findings, we expected some intrinsic differences between sets, and a general speed up over blocks, and indeed these effects were in evidence: For Blocks, $F1$ (2, 20) = 9.98, MSE = 2907, p = 0.001; for Set, $F1$ (3, 30) = 26.09, MSE = 1182, $p < .0001$. Two complex interactions, the Context × Set × Sequence interaction [$F1$ (3, 30) = 4.80, MSE = 577, p = 0.0076, and the Context × Set × Blocks interaction [$F1$ (6 ,60) =2.39 MSE = 673, $p < .05$] were significant as well but these may indicate only that spurious context effects for particular sets were not sustained across sequences or blocks. More important for our purpose, Context did not interact with Block: $F1$ (2, 20) = 2.37, MSE = 791, p = .1188. Thus, there was no evidence of attunement to homogeneous onsets as a function of protracted exposure to the materials. No other effects were significant.

The results of the by-items analysis (which collapsed Set and Sequence) showed only a significant main effect of Block: $F2$ (2, 30) = 46.62, MSE = 207, $p < .0001$, reflecting the

general acceleration over exposures to sets, but, consistent with the participants analysis, no effect of Context (F2 < 1) and no interaction of Context and Block.

Discussion

The goal of Experiment 1 was to replicate the null form preparation benefit for disyllable onsets reported by Chen et al. (2002, Experiment 5), using the same first syllables and design as in the following monosyllable experiments. The materials were simplified and disambiguated to optimize sensitivity to the onset preparation effect if it existed. The nonsignificant 3 ms overall effect provides no evidence of a preparation benefit for the onsets of these disyllables. Importantly, there was no variation in the preparation effect over the three blocks, and thus no evidence that experience with the procedure or with the items promoted form preparation. The complex interactions we observed in the participants analysis may merely reflect equilibration of the data as a function of counterbalancing.

Having replicated the null benefit of shared onsets in Mandarin form preparation, we next wanted to test the same effect with monosyllable targets. The null finding for onset preparation could be specific to complex words for which, according to our theory, Mandarin speakers must keep track of the order of syllables and select them in sequence. That is, the need to sequence syllables may keep Mandarin speakers focused at this level until after syllable selection. If so, preparation of onsets could emerge when only single syllables are in question. Alternatively, following the *proximate units principle,* Mandarin speakers rely comprehensively on syllables as primary phonological planning units. In that case, planning stays at the level of syllables, and the segmental ingredients of those syllables may not be deployed until the point when the speech plan is implemented.

There was no evidence of a benefit due to mere repetition of the same onset in homogeneous blocks. If such an execution effect occurred, it would be more evident in monosyllable production for the same reason that onset preparation might occur for monosyllables – production is limited to individual syllables and so does not contend with the complexities of sequencing that disyllables entail. Thus in the next experiments, there could be a slight benefit of homogeneity even if onsets are not planned.

Experiment 2: Preparing Monosyllables Extracted from Compounds

In this experiment, participants produced the same core syllables that were the first syllables of disyllable targets in Experiment 1, but as monosyllables. To implement this, we identified disyllabic compound words in which the core syllables appeared as the first or second constituent (see Appendix). We then used the other syllables of these compounds as prompts and the core syllables as targets in the form preparation procedure. English equivalents would be to use *arm* as forward prompt for *chair* or *man* as backward prompt for *chair*.

If word complexity, specifically the presence of more than one syllable, is what prevented onset preparation in Experiment 1, we should now observe the effect. If, however, what prevents onset preparation is that Mandarin speakers plan phonologically via proximate syllable units and not segments, little or no benefit will be observed.

Method

The design and procedure were the same as in Experiment 1. Only the materials differed. Each core syllable was paired with an associate that comprised the first or second syllable of an existing compound. Seven items were forward compounds, seven were backward, and two were bidirectional (see General Method and Appendix Table 2). The core syllables were always the targets, and the other constituents were prompts.

Results

As Table 2 shows, there was no benefit of sharing onsets. The overall error rate was 2% in the homogeneous condition and 1% in the heterogeneous condition. The overall onset preparation effect, −6 ms, was in the wrong direction but not reliably so: $F1$ (1, 10) = 2.59, MSE = 1035, p = 0.14. The context effects varied significantly across sets [Context × Set: $F1$ (3, 30) = 5.25, MSE = 616, p = 0.005], but this reflects only that the negative outcomes were larger in some sets. As in Experiment 1, Context did not interact with Block: $F1$ (2, 20) = 2.29, MSE = 424, p = .13.

As before, the Block and Set main effects were both significant: $F1$ (2, 20) = 11.46, MSE = 2513, p = 0.0005 and $F1$ (3, 30) = 36.72, MSE = 2438, p < .0001, respectively. People were faster overall if they started with a heterogeneous block [Sequence main effect: $F1$ (1, 10) = 5.75, MSE = 25627, p = .04], and the cost of preparation was associated with the slower homogeneous-heterogeneous sequence [Context × Sequence interaction: $F1$ (1, 10) = 8.12, MSE = 1035, p = 0.02], suggesting that the cost is an artifact of this asymmetry. The latter interaction was reduced over Blocks [Context × Sequence × Block interaction: $F1$ (2, 20) = 7.71, MSE = 424, p = 0.003], and was somewhat uneven over Sets [Context × Sequence × Set interaction: $F1$ (3, 30) = 2.91, MSE = 616, p = 0.05]. None of this qualifies our overall conclusion that there is no evidence of onset preparation.

Finally, the simpler items analysis showed only a significant main effect of Block: $F2$ (2, 30) = 75.08, MSE = 128, p < .0001.

Discussion

The results of Experiment 2 do not support the prediction that onset preparation would occur for monosyllable targets. The effect was in the wrong direction, though not significantly so, and the significant interactions with sequence suggest that the tendency may be artifactual. As noted, some benefit of homogeneity due to chronic activation of onsets or to lower level execution processes might occur even without preparation, but there was no evidence for this. Therefore, we conclude that there is no benefit for a Mandarin speaker in knowing the onset of a set of simple monosyllable targets in the form preparation procedure.

There is one possible concern with the materials in Experiment 2. We used disassembled compounds, with one constituent serving as a prompt and the other as response. This procedure ensured that the cues were strong reminders of the targets. However, the procedure may have induced participants to conceive of the prompts and responses as compounds rather than as discrete items, and thus to focus on the meaning of the compounds or to use the first syllables of forward compounds as the basis of planning. For backward and bidirectional compounds, there may have been some additional conflict about the order of syllables. Inspection of the items showed that the negative tendency was associated with the two bidirectional (−31 ms) and seven backward items (−5 ms). The seven forward compounds produced a literal null result (0 ms). Because our goal was to evaluate onset preparation under the most transparent and optimal conditions, we changed the cues in Experiment 3 so that prompts and targets were unambiguously mono-morphemes.

Experiment 3: Monosyllabic Monomorphemes

Participants produced the same targets as in Experiment 2 but the associative prompts were changed such that they did not form compound words with the targets. Instead the response words were all instantiated as mono-morphemes. If participants in Experiment 2 had failed to focus as intended on the response syllables, we did not expect that to happen in Experiment 3, and so if Mandarin speakers are disposed to prepare onsets, they should now do so.

Method

The design and procedure were the same as in Experiment 2. Only the associative prompts were changed. Each target was cued by a monosyllabic noncompound associate. See Appendix Table 3 for the full listing of materials.

Results

As Table 2 shows, the overall onset preparation effect, 2 ms, did not approach significance, *F1* (1, 10) < 1, *MSE* = 806. The error rate was equivalent (3%) in the homogeneous and heterogeneous conditions. The preparation effects did not vary across sets [Context × Set: *F1* (3, 30) < 1, *MSE* = 577], although the pattern did fluctuate a bit over Blocks [Context × Set × Block interaction: *F1* (6, 60) = 2.30, *MSE* = 441, *p* < .05].

Again, the Block and Set main effects were both significant: *F1* (2, 20) = 51.35, *MSE* = 947, *p* < 0.0001 and *F1* (3, 30) = 116.78, *MSE* = 1414, *p* < .0001. Context did not interact with Block: *F1* (2, 20) < 1, *MSE* = 714, so there is no evidence of emergence of a preparation benefit. As in Experiment 2, although there was no main effect of sequence, the effect of Context was negative in the homogeneous-heterogeneous sequence (−16 ms) but positive in the heterogeneous-homogeneous order (20 ms) [Context × Sequence interaction: *F1* (1, 10) = 28.25, *MSE* = 806, *p* = 0.0003]. However, this pattern was mitigated over Blocks [Context × Sequence × Block interaction: *F1* (2, 20) = 8.82, *MSE* = 714, *p* = 0.002]. We offer an explanation of these interactions following the report of Experiment 4 below.

As was the case for Experiment 2, the results of the by-item analysis yielded only a significant main effect of Block: *F2* (2, 30) = 165.76, *MSE* = 98, *p* < .0001.

Discussion

Again we observed no benefit of homogeneous onsets. This suggests that whether prompts and targets form compounds (Experiment 2) or not, Mandarin speakers address targets in the same way. The absence of an onset preparation benefit suggests that this way is different than in European languages: Mandarin speakers plan syllables and do not engage onset information even though this would potentially provide a benefit in homogeneous contexts. Before accepting this conclusion, we considered one more reservation, namely that despite our efforts to guide speakers to focus on monosyllable targets, they may have nonetheless treated them as parts of compounds.

Although the materials in Experiment 3 were monosyllabic prompt-response associates, Mandarin is highly productive in forming compounds (Packard, 2000). Thus, it is possible that participants intentionally or implicitly treated the prompt-response pairs as novel compounds. In previous work on novel compounds, we have found that compounding occurs much more readily in Mandarin than in English (O'Seaghdha et al., 2003). Also there are some lingering questions raised by the sequence interactions. In both Experiments 2 and 3, a negative tendency was associated with the homogeneous-heterogeneous sequence. Could this be diagnostic of a strategy that blocked the expression of the onset preparation benefit? To address these concerns, in Experiment 4 we used nonlinguistic cues to remove any possibility that our participants could spontaneously compound the prompts and targets and thus inadvertently block their ability to prepare the onsets of monosyllables.

Experiment 4: Nonlexical Prompts

The targets were again exactly the same but in this experiment the associative cues were nonalphanumeric keyboard symbols. We reasoned that use of symbols ensured that participants

prepared only the associated monosyllable targets and could not be influenced by existing (Experiment 3) or induced (Experiment 4) compounds.

Method

The method was the same as in the previous experiments. The prompts were non-alphanumeric keyboard symbols such as #, @, &, as shown in Appendix Table 4. These symbols were scaled to match the size of the characters. Participants had no difficulty learning the arbitrary linkages of four of these symbols to the responses in a set, and the results show that latencies were comparable to those in the preceding experiments.

Results

As in the previous experiments, the error rate, 4% in the homogeneous condition and 3% in the heterogeneous condition, did not vary systematically (see Table 2). The overall onset preparation effect of 4 ms was not significant: $F1$ (1, 10) < 1, MSE = 2542. Again, the preparation effect varied significantly across sets [Context × Set interaction: $F1$ (3, 30) = 4.50, MSE = 1023, p = 0.01]. However, none of the sets showed a substantial positive benefit.

The Block and Set effects were both significant: $F1$ (2, 20) = 11.41, MSE = 4455, p = 0.0005 and $F1$ (3, 30) = 7.44, MSE = 2890, p = .0007. However, unlike in Experiments 2 and 3, Context did not interact with Sequence, $F1$ (1, 10) < 1, MSE = 2542, and no other effects were significant.

The results of the by-item analysis showed only a significant main effect of Block: $F2$ (2, 30) = 47.21, MSE = 359, p < .0001. There was no effect of Context, and Block did not modulate the Context effect.

Combined Analysis of Monosyllable Experiments

Our theory links the absence of an onset preparation benefit in Mandarin to the status of syllables as selectable units linked to word/lemma representations. Not only do Mandarin speakers use syllable units to plan and regulate speech, but this language property appears to block them from engaging the shared onsets of those syllables even under conditions such as those of Experiments 2, 3, and especially 4, where intuition suggests they are "there for the taking".

Taken together, the monosyllable experiments indicate that the status of prompt response pairs as compounds (Experiment 2), potential compounds (Experiment 3), or as isolated monosyllables (Experiment 4) has little influence on performance. To confirm this point, to increase power, and to consolidate the interpretation of other patterns in the data, we performed a combined analysis of Experiments 2-4 with Experiment as a between groups factor.

In this analysis, there was no evidence whatever of a Context effect, $F1$ (1, 30) < 1. However, consistent with Experiments 2 and 3 individually, sequence made a difference [Context × Sequence interaction: $F1$ (1, 30) = 16.26, MSE = 1461, p < .001]. The pattern of interaction is that when the homogeneous blocks are first, there tends to be a negative context outcome, but when they are second the pattern is facilitatory. It would be problematic for our theory if indeed the null result of Context reflected a mixture of positive and negative tendencies in the two different sequences. However, meaningful inhibitory effects have never been reported in the form preparation task and we did not prognosticate them here. Fortunately, the pattern can be accounted for as artifactual.

There seems to be an anti-preparation effect when the homogeneous blocks are first, but this is due only to the relative slowness of the first blocks in all conditions. Conversely, there is

seeming facilitation when the heterogeneous blocks are first for the same reason. The sequence disparity therefore diminishes over Blocks [Context × Sequence × Block interaction: $F1$ (2, 60) = 11.61, MSE = 1193, p <.0001, also seen individually in the analyses of Experiments 2 and 3]. The sequence effects and interactions, though not equally evident in Experiment 4, did not differ statistically across experiments. Thus, we conclude, that there is no form preparation of onsets in any of the monosyllable experiments and that the seeming variation in the preparation effect with sequence is an artifact of the confounding of sequence with early task experience.

The Context effect did vary significantly across sets in the combined analysis [Context × Set interaction: $F1$ (3, 90) = 5.51, MSE = 739, p = .0016]. Overall, the /m/ set was inhibitory (−13 ms) whereas the other three sets showed small positive net effects. Items in the /m/ set, however, were not consistently inhibitory. Furthermore, exclusion of the /m/ set did not produce a significant context effect. We report set effects for consistency with a convention in some of the relevant form preparation literature, but we see no theoretical basis for focusing on set effects per se. Rather, we suggest that the appearance of negativity for the /m/ set may be a product of systematic item tendencies in a difference distribution that is centered near zero.

Sets also varied considerably in overall ease of production, $F1$ (3, 90) = 101, MSE = 2247, p < .0001, and this varied somewhat across experiments [Set × Experiment interaction: $F1$ (6, 90) = 10.87, MSE = 2247, p < .0001], perhaps because of variation in the affinities of prompt types and response sets. Finally, there were main effects of Experiment, $F1$ (2, 30) = 6.98, MSE = 37131, p < .005, and Block, $F1$ (2, 60) = 46.66, MSE = 2638, p < .0001.

Discussion

Once again, we found no preparation effect in Experiment 4. This experiment is particularly convincing because the use of nonlinguistic prompts required participants to process the targets as monosyllables and left no basis for compounding. Although we used nonalphanumeric symbols as prompts, neither latencies nor error rates indicated that the task was more difficult than with lexical prompts. The combined analysis across Experiments 2-4 likewise showed no evidence of an onset preparation benefit. We conclude that Mandarin speakers do not prepare monosyllable onsets under conditions where speakers of European languages clearly do (Alario et al., 2007; Meyer, 1991; Roelofs, 2006; Santiago, 2000).

Although the case for our claim that Mandarin speakers prepare syllables and not segments appears to hold for monosyllables as well as for disyllables, two considerations could weaken this claim. First, Mandarin monosyllables are simpler than those used in previous demonstrations such as the Dutch experiments of Meyer (1991). To provide new positive evidence of onset benefits in English with materials similar in structure to our Mandarin monosyllable targets, we used monosyllable consonant-vowel targets in two English language experiments. In Experiment 5, we used monosyllabic lexical prompts with simple consonant-vowel response words in a design similar to that of Experiment 3.

A second alternative explanation for failure to prepare Mandarin onsets is that preparation may depend on alphabetic cuing. In languages with alphabetic writing systems, the structural similarity of target beginnings may be made more salient when they share the same visible first letter during learning, and at least for English, form preparation may be disabled when they do not (Damian & Bowers, 2003; but see Roelofs, 2006). However, the absence of alphabetic orthography is an unlikely explanation for failure to prepare onsets in Mandarin. Roelofs (2006) has shown that phonology trumps letter consistency in most circumstances in Dutch, and that removing alphabetic cues through the use of picture prompts actually yields larger not smaller preparation effects. Alario et al. (2007) also used picture prompts in order to circumvent orthography and found clear onset preparation benefits in French regardless of letter

consistency. To extend these previous findings to simple consonant-vowel monosyllables, in Experiment 6 we used nonlexical picture prompts to elicit production of the same targets as in Experiment 5.

Experiment 5: English Monosyllables Prompted by Associates

The purpose of Experiment 5 was to test for form preparation of simple consonant vowel English monosyllables, similar in structure to the type of Mandarin syllable we used in Experiments 1-4. We predicted a significant preparation benefit.

Method

Participants—Sixteen Lehigh University undergraduates enrolled in introductory psychology participated for a research experience credit. All were native English speakers.

Design—The design was equivalent to that of Mandarin Experiments 1-4. Eight participants served in each context sequence.

Materials—Four homogeneous sets of four targets were selected. Because of the English restriction on monosyllable CV words with short vowels, we searched for monosyllabic words beginning with a particular consonant and continuing with a long vowel or diphthong. Only sets for which four different continuations were available could be used. In addition the target words were limited to nouns that could be prompted pictorially in Experiment 6. These constraints narrowed down the viable sets to the four we used: /d/, /p/, /r/, and /s/. All of the words were common except *roux* which was unfamiliar to unculinary participants but was defined for them prior to the experiment (see Procedure). The heterogeneous conditions were arranged such that each word began with a different consonant and contained a different vowel. The full set of materials with their associated prompts is shown in Appendix Table 5. The prompts were also monosyllables and had a transparent connection to the response words.

Procedure—The procedure was structurally the same as in the Mandarin experiments but differed in some details. Participants were run individually on a PC running E-Prime 1.0 software. Naming times were recorded using a microphone voice-key interfaced to the E-Prime program, and an additional audio record was created for error coding using a Nomad Jukebox 3. Participants were first introduced to the response words in random order to ensure that their intended meanings were known and that they were produced with the intended pronunciation. Each word was shown on the screen together with an example sentence (e.g., DAY as in: It is a nice day). Participants read each word aloud to enable the experimenter to check the pronunciation and correct it if necessary. In addition, we provided definitions for *roux* ("a cooked mixture of flour and fat used as a thickening agent in a soup or sauce") because many participants were unfamiliar with this word, and *sow* ("an adult female pig") because pilot participants tended to pronounce it as in the verb *to sow*.

Participants were next instructed on the testing procedure and given practice with one homogeneous (*sting bee; port bay; ghost boo; leave bye)* and one heterogeneous set (*mud goo; scale weigh; lock key; clothes tie*).

As in the Mandarin experiments, each session comprised a learning phase and a testing phase. During the learning phase, the four prompt-response pairs for each set were first shown on the screen. When the participant was ready to proceed, additional learning was provided by presenting each prompt alone at center screen for 1000 ms followed by the centered prompt-response pairing for an additional 1000 ms. During the testing phase, each trial began with the presentation of an asterisk in the center of the screen for 100 ms followed by a 250 ms warning tone. After a 250 ms delay, the prompt appeared in the center of the screen for 150 ms.

Participants were instructed to say the response words aloud, as quickly and accurately as possible. If there was no response within 1000 ms of prompt offset, a timeout was signaled by a 250 ms tone. The next trial began 1250 ms after the onset of the response or the timeout signal.

Results and Discussion

After elimination of responses occurring in the first 200 ms from prompt onset, timeouts (> 1000 ms from prompt offset), and miscellaneous errors, mean response times were calculated for the valid responses. Table 3 shows that there was an overall 12 ms preparation benefit.

Analysis of variance on the latency data showed the expected significant effect of context, $F1$ (1, 14) = 7.18, MSE = 1993, p = .018. As in the Mandarin experiments, there was some intrinsic variation associated with sets, $F1$ (3, 42) = 5.03, MSE = 4572, p = .005. Also, latencies decreased over Blocks (699 ms, 664 ms, and 654 ms), $F1$ (2, 28) = 26.09, MSE = 2798, $p <$ .001, but there was no modulation of the Context effect over Blocks.

Analysis by items showed a significant effect of Context, $F2$ (1, 15) = 6.17, MSE = 578, p = .025, and also of Block, $F2$ (2, 30) = 64.40, MSE = 270, $p <$ .001.

Experiment 5 shows a form preparation benefit in English for simple consonant-vowel monosyllables, whereas no such effect was observed in the comparable Mandarin experiments.

Experiment 6: English Monosyllables Prompted by Pictures

Although the response words were not visible during the testing phase of Experiment 5, the prompts were alphabetic and so could have influenced participants to think of the responses as decomposable into segments. In the Mandarin monosyllable experiments, in contrast, the prompts were single characters in Experiments 2 and 3 and nonlexical symbols in Experiment 4. In Experiment 6, following Alario et al. (2007) and Roelofs (2006), we used another type of nonalphabetic prompt, pictures. As noted by Alario et al., in addition to circumventing the need for alphabetic prompts, the use of pictures reduces the role of memory in the form preparation task and thereby increases its validity as a window on production.

Method

Participants—Sixteen new participants were recruited from the same pool as in Experiment 5. Half of them served in each sequence.

Materials—The response words were the same as in Experiment 5. A transparently related picture prompt was selected for each response word (see Table 4 for examples).

The picture prompts directly evoked the target words. They were located through a Google Image search of the target word or a close synonym of the word. Pictures were photographic images and were selected based on simplicity and clarity. All pictures were 300×300 pixels, resized to approximately 7.5cm horizontally and 6cm vertically, and converted from .jpeg to .bmp format for compatibility with the E-Prime program.

Procedure—Except for the use of picture vs lexical prompts and a small difference in cue contingency, the procedure was the same as in Experiment 5, including familiarization, learning, and testing. Because the pictures were transparently linked to the responses, we found it more natural to leave them on the screen during the response interval. Thus pictures remained on the screen until the participant responded or until the trial timed out after 1150 ms. Participants found it easy and enjoyable to learn and produce the responses with picture prompts.

Results and Discussion

The data were coded and processed in the same way as for Experiment 5 (see Table 3). The Context effect of 16 ms is significant, $F1$ (1, 14) = 6.93, MSE = 3678, p = .02. Sets also differed, $F1$ (3, 42) = 7.97, MSE = 3006, $p < .001$, and the context effect varied a bit with set [Context by Set interaction: $F1$ (3, 42) = 2.85, MSE = 1012, p = .05. Finally, there was an effect of Blocks, $F1$ (2, 28) = 3.75, MSE = 2114, p = .04, which reflected not the usual acceleration but rather a slight slowing (first 560, second 564, third 575). We assume that this reflects the relative ease of the task which left little room for improvement early in the experiment, whereas a slight entropy eventuated in the third block. No other effects or interactions were significant.

The analysis by items showed a significant Context effect, $F2$ (1,15) = 15.33, MSE = 429, p = .001. The Block effect was also significant, $F2$ (2, 30) = 10.04, MSE = 203, $p < .001$. Again there was no interaction between these two variables.

A combined analysis of Experiments 5 and 6 confirms that the priming effect is robust, $F1$ (1, 28) = 13.76, MSE = 2835, p = .001, $F2$ (1, 30) = 19.70, MSE = 504, $p < .001$, $min\ F'$ (1, 56) = 8.1, p = .006. Both the individual analyses of Experiments 5 and 6 and the combined analysis thus support our conclusion that onset preparation is normative in English, even with these very simple words. We interpret the modest effect size as due to the simplicity and familiarity of most of the words we used. Preparation benefits tend to be larger with less familiar, low frequency, and more complex items (see Meyer, 1990, 1991) because there is more room for gain. Taken together, Experiments 5 and 6 show clear benefits of onset preparation for English monosyllables, contrasting with the failure to find any benefit for comparable items in Mandarin.

Experiment 7: Syllable vs Onset Preparation

Our agenda hinges on the finding of a first atonal syllable but not an onset preparation benefit in Mandarin. By atonal syllable we mean the syllable with tone unspecified, so that speakers do not know exactly how it will be articulated. Whereas tonal syllable preparation yields a robust benefit, the atonal effect is quite small, 12 ms in Chen et al. (2002, Experiment 5). Thus it is crucial though challenging to confirm the reality of this effect and to contrast it with a null or near null onset benefit. To most effectively meet these goals, we devised a larger form-preparation experiment in which both effects were assessed on the same target words. The full design, shown in Appendix 1 Table 6, used three sets of three syllables beginning with the same onsets, /f/, /k/, and /p/. These three replications of the syllable conditions are shown in the left side of the table. In the onset conditions, shown in the right side of the table, these same items were recombined such that each homogeneous set comprised three different syllables, but still with the same /f/, /k/ and /p/ onsets. Participants were tested in three groups, each receiving one syllabic and one onset matrix, so that over all participants the identical words were tested in the syllable and onset conditions but repetition of words across these conditions for individual participants was minimized. Based on our theory and the previous findings of Chen et al. (2002), we predicted a modest benefit of atonal syllable priming. As in Experiments 1-4, we predicted little or no benefit of shared onsets.

This experiment followed the general method of Experiments 1-4 but with some modifications as specified here.

Method

Participants—This experiment used 36 undergraduate students from National Cheng Kung University, in three groups of twelve. They were all native Mandarin Chinese speakers and were paid for participation.

Apparatus—The experiment was programmed in E-Prime 1.0 on a PC.

Materials—The materials for the atonal syllable conditions involved three sub-designs, each comprising three sets of three disyllables, each homogeneous set defined by a different first atonal syllable and one of three onsets, /f/, /k/, or /p/. The Homogeneous sets for the first group shared the syllables *fei, ke,* or *pi*, for the second group, *fen, kan,* or *pao*, and for the third group, *fan, ku,* or *pu*. The Heterogeneous sets comprised words beginning with one of each of the respective syllables. See Appendix Table 6 for full details. To form the Homogeneous onset conditions, the syllabically grouped words were recombined into sets sharing only onsets. For example, the three /f/ sets comprised the first syllables *fei, fen,* and *fan,* but with different lexical items in each set. Thus, the three subdesigns were balanced so that each word occurred as a member of the syllable condition and as a member of the onset condition for different participants. One item was necessarily shared between homogeneous syllable and onset sets for each group of participants, and one heterogeneous set was repeated. Note that because participants were tested in two matrices rather than one, we used three items per set, whereas Chen et al (2002) and Experiments 1-6 of this paper used four item sets. However, the use of three-item sets is common in the form preparation literature.

Design—As explained above, this experiment used three counterbalancing subdesigns, with each Group of 12 participants receiving two Types of materials, syllabic and segmental. Type Order was counterbalanced with half the participants receiving conditions involving atonal syllables first and half receiving the segmental onset conditions first. The primary manipulation as in all of the experiments was Context, and the Sequence of contexts was counterbalanced across participants. Half of the participants received homogeneous contexts first and half received heterogeneous contexts first, making four combinations of Type Order and context Sequence, with the same sequence applied to syllabic and segmental types for an individual participant. Each of the three Sets in a context was presented in random order before switching to the other context. In a presentation block, the three prompt-response items in a set were presented in random order four times (Repetitions) and the entire Block procedure was replicated three times before moving to the next set. Altogether, each participant received a total of 432 trials (2 types × 2 contexts × 3 sets × 3 items × 4 repetitions × 3 blocks). Group, Sequence, and Type Order, were between groups factors and all other factors were within-subjects.

Procedure—We used the same associative procedure as in the preceding experiments. During the learning phase, the participants saw the three prompt-response pairs of a set on the screen. The procedure was the same as in Experiments 1-4 except that three heterogeneous practice sets rather than just one were given before the experiment began. It took about one hour to complete a session.

Results

To provide a full account of the findings, the results are shown comprehensively in Table 5 by group, by set, and by block. Following the block results, the combined results for blocks 2 and 3, excluding block 1, are also given with reference to the separate analysis of those data. The conventional presentation by sets shows evidence of syllable preparation for two of the three sets. The preparation effects in the atonal syllable condition were 8, 12, and 7 ms for the three groups. In the onset condition they were 2, −1 and 5 ms. Thus the overall effect was 9 ms in the atonal syllable condition and 2 ms in the onset condition. The Block view (discussed further below) shows no context effect in the first syllable block (3 ms) or in any of the onset condition blocks, but an increasing context effect in syllable blocks 2 (9 ms) and 3 (15 ms).

Overall Analysis—For ease of presentation of this more complex design we report analyses of variance by participants and by items together. Type, Context, Set, and Replication were within-participants variables, and Type Order and Sequence were between-groups factors. In the by-items analysis, following Chen et al., (2002), we collapsed Set and the between-groups factors and evaluated effects of Type, Context and Block.

The main effect of Context was marginal by subjects, $F1$ (1,24) = 3.15, MSE = 3236, $p = .09$, but significant by items, $F2$ (1, 26) = 15.52, MSE = 185, p = .0005. The key prediction, a preparation benefit for syllables but not for onsets was marginally supported by subjects but again significant by items [Context by Type interaction: $F1$ (1,24) = 3.75, MSE = 1079, p = .06; $F2$ (1, 26) = 7.54, MSE = 154, p = .01]. There were also two complex interactions involving Context, Type, Sequence and Group, $F1$ (2, 24) = 3.58, MSE = 1079, p = .04, and Context, Type, Order, Sequence and Group, $F1$ (2,24) = 4.14, MSE = 1079, p = .03.

Other noteworthy effects include main effects of Set, $F1$ (2,48) = 22.49, MSE = 1532, p = .0001, and Block $[F1$ (2,48) = 59.91, MSE = 1056, p = .0001; $F2$ (2, 52) = 64, MSE = 238, p < .0001], the latter indicating the typical speedup of responses with extended exposure to the sets. The Block effect was involved in a complex interaction with Order, Sequence, and Group, $F1$ (2, 48) = 3.53, p = .01, suggesting complex variability in acceleration over blocks that is relevant to our decision to analyze the data without the first block below. The main effect of Set cuts across different items and so does not have an obvious interpretation. It was modulated by Group as well as by the combination of Order and Sequence, though not by Type. Although the items were the same in both conditions, there was a 7 ms overall advantage of syllable conditions, perhaps as a by-product of facilitation in that condition. This main effect of Type was not significant by subjects, but was significant by items, $F2$ (1, 26) = 5.32, MSE = 599, p = .02. None of the between groups factors was significant.

The crucial interaction whereby a preparation benefit was predicted for syllables but not for onsets, was supported, but because the effect in the analysis by subjects is marginal the statistical verdict is not entirely compelling. We therefore considered whether our design might account for the slightly weaker atonal syllable benefit (9 ms) in this experiment relative to the 12 ms effect in the experiment of Chen et al (2002). One consideration is that we used three-item rather than four-item sets perhaps moving performance toward ceiling and reducing the space for expression of the small preparation benefit. Another is that our design allowed repetition of one heterogeneous set across types in each group. This could speed responses in the later heterogeneous condition containing the repeated set, again reducing the size of the preparation benefit. Finally, as noted in the introduction, earlier blocks, especially at the beginning of the experiment, but also after a switch of conditions, may be noisier in general.

Inspection of the data showed that syllable preparation was absent in the first block when syllables were tested first, regardless of the order of homogeneous and heterogeneous conditions, whereas a syllable preparation benefit was evident in all other circumstances. This suggests that participants picked up on the benefit of shared syllables immediately when they featured in the second half of the experiment, but for some reason were not so adept in the beginning. In contrast, the onset preparation outcomes were equally distributed between positive and negative values throughout, and there was no evidence of attunement to the presence of shared onsets after extended experience. We therefore conducted a secondary analysis excluding Block 1 data.

Analysis of Blocks 2 and 3—This analysis is identical to the overall analysis but with the first block regardless of Type Order or Sequence removed. The effect size for syllable preparation increased to 12 ms, exactly the same as in Chen et al. (2002). The effect size estimate for onset preparation was unchanged at 2 ms (see Table 5).

There was an overall effect of Context, marginal by participants [$F1$ (1, 24) = 4.17, MSE = 2480, $p < .06$; $F2$ (1, 26) = 21.06, MSE = 138, $p < .0001$]. More important, this effect was significantly modulated by type of form *p*reparation: $F1$(1, 24) = 5.41, MSE = 1056 , $p < .05$; $F2$ (1, 26) = 9.66, MSE = 142, $p < .005$; *min F'* (1, 45) = 3.47, $p = .069$. Thus, when Block 1 was excluded, the theoretically crucial restriction of form preparation to the syllabic condition was supported.

There were also significant main effects of Block [$F1$ (1, 24) = 22.90, MSE = 865 , $p < .0001$; $F2$ (1, 26) = 18.03, MSE = 275, $p < .01$], and Set, $F1$ (2, 48) = 21.70, MSE = 1244, $p < .0001$, as well as a number of smaller, theoretically uninteresting interactions with these factors involving blocks, groups and the order of conditions. Type was significant by items only, $F2$ (1, 26) = 8.26, MSE = 464, $p < .01$.

Discussion

This experiment establishes the contrast between significant atonal syllable priming and the null benefit of shared onsets in a powerful within subjects and items design. The findings provide strong corroboration of the previous across experiment results of Chen et al. (2002). The use of the same items in both the syllable and onset conditions is especially important because it removes any concern that the preparation difference could be due to variation in the items themselves or in the effectiveness of the cues used to elicit them. Moreover, our design provides three independent and largely consistent replications through our three counterbalancing groups. The null onset preparation effect contrasts not only with the significant atonal syllable benefit within the experiment but also with significant shared onset benefits in Dutch (e.g., Meyer, 1991; Roelofs, 2006), English (Damian & Bowers, 2003), and French (Alario et al., 2007), and in English Experiments 5 and 6 above.

The data tell the same story as a whole and in the secondary analysis that excluded all Block 1 data. However, the syllable preparation benefit was greater in the latter, and the statistical backing of our hypothesis is correspondingly more convincing. Specifically, the interaction of Type and Context was only marginally significant by subjects though significant by items in the omnibus analysis, but was significant in both respects in the reduced analysis. Although it would be simpler if the interaction were more robustly significant in the omnibus analysis, the pattern is quite clear throughout the data, with the one exception of the syllable preparation benefit when this condition is first encountered early in the experimental session. Because nothing in our theory requires that the syllable preparation benefit be equally evident in the early stages of performing an unfamiliar task as in later ones, we conclude that the benefit in Blocks 2 and 3 is real.

In contrast to the atonal syllable preparation benefit, a benefit of onset homogeneity was nowhere in evidence. These further replications of the null benefit of shared onsets with completely different materials strongly corroborate our contention that Mandarin speakers do not show benefits of onset homogeneity under conditions where speakers of Indo-European languages clearly do.

General Discussion

Previous form preparation research with European languages has shown clear evidence of advance preparation of word onsets for monosyllables (Meyer, 1991) and disyllables (Alario et al., 2007; Damian & Bowers, 2003; Meyer, 1990; Roelofs, 2006). This benefit is not a function of alphabetic processing because it is also observed robustly in a picture naming task (Alario et al., 2007; Roelofs, 2006; and our Experiment 6). In contrast, Chen et al. (2002) found no benefit of shared disyllable onsets in Mandarin, but a clear though intrinsically small preparation benefit for shared atonal first syllables. Our findings a) replicate the Chen et al.

(2002) null result for onset preparation with refined disyllable materials; b) expand the empirical and theoretical purchase of this finding by demonstrating the same lack of onset preparation for simple monosyllables in Mandarin; c) contrast this with a clear onset preparation benefit in English with comparably simple monosyllabic words; and d) replicate and consolidate both the positive atonal syllable benefit and the null onset preparation effects in a powerful within subjects and items design. Taken together, these findings are consistent with the view that the original or *proximate* phonological encoding units in Mandarin are syllables whereas in English and related languages they are phonemic segments.

Our findings suggest not only that phonological syllables play a distinctive, primary role in phonological encoding in Mandarin and other Chinese languages, but that they also prevent access to the seemingly obvious benefit of homogeneous over heterogeneous onsets. We address these two linked findings in turn, because they provide distinct insights.

Atonal syllable preparation—In our theory, atonal syllables are the proximate units of Mandarin, units that despite their tonal vagueness, speakers can prepare for, think about, and further specify. We chose to examine atonal rather than tonal syllable preparation (Experiment 7), despite the inherent weakness of the effect, because it is much more theoretically revealing. Whereas almost any theory would predict a benefit of knowing exactly what the first syllable of a word will be, the nature of the benefit is not easy to specify. It could include effects of initial retrieval of phonological information as well as later pre-articulatory processes. In contrast, atonal syllables are not specific enough to guide articulation, and so the preparation benefit can not be prearticulatory. Thus, the atonal syllable benefit supports our hypothesis that atonal syllables are proximate units that may be employed in the planning of speech. It follows that form preparation may occur at a relatively abstract, prearticulatory, phonological level.

In contrast to segments which are not effectively planned if the set members are not unanimous (e.g., two /d/ onsets and a /t/ onset; Roelofs, 1999), atonal syllables do provide a preparation benefit. It is useful to compare the two cases, atonal syllables with a different tone for every item, and segments varying in one phonetic feature of just one of the items in a set, because on the surface the difficulty is greater in the syllable case. What makes it possible to benefit from the atonal syllable and not the near-unanimous segments? We propose that the difference is precisely that the atonal syllable is a proximate unit, whereas a set of less-than-unanimous segments does not point to a unique unit. The atonal syllable can be kept in mind until the tone is supplied, the unresolved segment does not have a coherent definition across the full set and so is not effectively addressed by the planning process.

No Onset Benefit in Mandarin—The negative result, the absence of an onset preparation benefit in Mandarin, and especially its absence for monosyllables, is both counterintuitive and very informative. It is surprising because of the many demonstrations of onset preparation benefits in other languages. Our Experiments 5 and 6 were specifically designed to assess the onset benefit with monosyllables of comparable simplicity to the Mandarin items and showed a robust effect. Why then did we not observe an onset preparation benefit in any of the Mandarin experiments?

As we have discussed, there are several ways in which an onset benefit could arise in Mandarin. First, one might expect some benefit due to the mere repetition of onsets in homogeneous sets. The fact that this did not occur suggests that form preparation is not an undiscriminating implicit process, but rather that it requires intentional orientation to the phonological component. Note that Meyer (1990) proposed that mere repetition accounted for little of the benefits in her original form preparation experiments. However, it was not possible to segregate effects of mere repetition of segments from positive effects of preparation of the same units in Meyer's Dutch experiments. The null effect of onsets in our Mandarin experiments shows that mere

repetition has no effect, at least under these experimental circumstances. This leaves the question of why participants did not recognize the potential benefit of shared onsets in the homogeneous contexts of our Mandarin experiments and prepare accordingly. We found no evidence of increasing attunement to onsets over blocks in the experiments.

Experiments 2-4 showed that the obliviousness to onset homogeneity in Mandarin is pervasive, applying to simple monosyllabic words as well as to polysyllables. In light of this outcome, we propose that what restricts onset preparation is that there is a requisite step of syllable encoding between word selection and segmental specification in Mandarin but not in European languages such as English (see Figure 1). Under this proposal, advance preparation in Mandarin is stationed at the syllable level, not the level of segments, and so there is no benefit of sharing onsets whether the set of words under consideration is disyllabic or monosyllabic.

What is Form Preparation?—Our findings suggest that the form preparation procedure is not just a tool that reveals the functional units of phonological encoding in various languages (e.g., Alario et al., 2007; Chen et al., 2002; Damian & Bowers, 2003; Kureta et al., 2006; Meyer, 1990, 1991; Roelofs, 2006; Santiago, 2000), but one that affords important insights into the nature of the representations involved in preparation as distinct from execution of speech. The atonal syllable benefit suggests that if the prepared units are selectable, they can be abstract relative to prearticulatory encoding. Because our theory says that segments are selected in Mandarin, just not during the initial access to proximate syllable units, the absence of the onset benefit in Mandarin suggests that selectability is not enough to ensure preparation. In addition, speakers must intentionally orient to the component in question. Our findings show that Mandarin speakers either cannot or simply did not orient to nonproximate word onsets.

A key question for future research is how Mandarin speakers represent prepared atonal syllables. One possibility, similar to what Meyer (1990) proposed for segmental preparation, is that they repeatedly mentally simulate the phonemic specification of the syllables in inner speech. However, for the segmental case, two findings suggest that prepared speech may not be featurally specified. First, as noted earlier, Roelofs (1999) showed that sets of whole phonemes but not sets of featurally similar phonemes show preparation benefits. Second, Oppenheim and Dell (2008) recently reported that silently rehearsed English speech, in contrast to overt speech, does not show an influence of featural similarity on phoneme errors. In our terms, form preparation and covert speech engage proximate but not subproximate representations. If this generalizes to syllables in Mandarin, form preparation of atonal syllables may entail speakers intentionally orienting to whole syllable units without impleting their segments. That is, knowing that upcoming items begin with a certain syllable is not necessarily implemented as rehearsal of the production of that syllable.

This question of how preparation occurs, and how it relates to the later processes of encoding is not well addressed in current models of word production which have focused primarily on immediate encoding. For example, in Dell's model, selection entails the provision of a jolt of activation to the selected unit that then reverberates through the network (Dell, 1986; Dell & O'Seaghdha, 1991). If such a momentary jolt were supplied to a Mandarin syllable or English onset in a form preparation block, it would need to be constantly replenished or reinstituted between productions to be effective throughout the execution of the block.

The WEAVER++ model addresses form preparation and other planning effects more explicitly through a suspend-resume mechanism (Roelofs, 1997; Levelt et al., 1999), whereby processes at different stages of encoding (such as phonological, phonetic) can be held in memory and reengaged at a later time. Whereas this proposal avoids the need for constant reselection, it also lacks psychological plausibility. If form preparation can involve incompletely specified components such as syllables without a determinate tone, what is prepared and what is later

produced are not the same thing, and so the latter is unlikely to be simply the resumption of the former. Rather we suggest that preparation involves a different process that is relevant to the later selection and production of a word, but is not simply buffered during the pre-production interval. This view is consistent with the separation of pre-motor and motor processes in physical action planning and execution (see e.g., Hickok and Poeppel, 2007; Eickhoff et al., 2009, for discussion of the corresponding language pathways), and allows for the maintenance of a stable representation of the shared element of a homogeneous set independently of the fluctuating activity in the production pathway. Thus, in our view, selectability is a primary qualification for proximate unit status, but this does not mean that units are selected in the same way in preparation and in actual word production. Clearly this is an area where more research, including theoretical, neuroscientific, and modeling work is needed.

Proximate Units and Prearticulatory Representations—A key to the ability of Mandarin speakers to prepare atonal syllables may be that atonal and tonal syllables are separated by only one attribute or feature. Whereas atonal syllables are proximate with respect to lexical retrieval, tonal syllables comprise the *penultimate* representation that precedes the actual articulation of speech. In our theory, the transition from atonal to tonal syllables involves linearization of each syllable's segments, and application of the appropriate tonal value (see Chen et al., 2002). However, although segmental speech errors show that segments are discretely represented during this transition, the small groups of retrieved segments for each syllable fully determine its assembled form. In that sense it is only the application of tone that distinguishes atonal proximate syllables from their penultimate tonal versions. What we want to suggest here is that this is what enables a speaker to easily "reverse engineer" atonal syllables in the form preparation task as well as in other contexts. It is easy for a Mandarin speaker to conceptualize a syllable shorn of its tone. Thus atonal syllables may be mentally represented though they cannot be spoken.

In contrast, the distance between proximate and penultimate representations in languages in which segments are proximate units is much greater. This is spelled out most clearly in the Levelt et al. theory where segments are assembled into *phonological words* that can vary considerably in complexity. A phonological word contains one stressed syllable but may include more than one morpheme (e,g., *the horse, horses, horsing,* are all phonological words). However, form preparation involves units smaller than lexical items and so it follows that any pronounceable word-initial fragment can also be an *ad hoc* phonological word. This leads to the question of how single consonant onsets are represented. Just as each syllable unit had a different tone in the syllable conditions of Experiment 7, each onset in Experiments 5 and 6 had a different vocalic continuation. Under those conditions, and because many consonants (including /d/ and /p/ in our materials) cannot be produced without some additional vocalic specification, it seems unlikely that participants could simulate the production of these onsets. Rather, it seems that the onsets must be represented abstractly, without rehearsing their production. In conditions where words share more than onsets (e.g., if they share a consonant and following vowel), the speaker can anticipate a phonological word fragment which may then move preparation to the prearticulatory level. Thus, the source of the preparation benefit for a consonant-vowel or any other longer, coherent fragment becomes ambiguous in the same way as for a tonal syllable in Mandarin: the benefit reflects prearticulatory preparation and may or may not also incorporate earlier stages.

If our conjecture that atonal syllables and isolated onsets are both prepared in a different and more abstract way than tonal syllables and multi-segmental fragments, respectively, is correct, it may be more than coincidence that the benefit of preparing the first of two parts of an English monosyllable in Experiment 5, and the first of two atonal syllables of a Mandarin disyllable (Blocks 2 and 3 of Experiment 7), is approximately the same.

Cross-linguistic generality—Our findings here, as well as the emerging consensus concerning the architectural difference between Chinese and Indo-European languages point to a general dilemma for theories of language production. It is becoming clear that theories of language production that are formulated in terms of specific units are limited in generality. For example, the Dutch flavored theory of Levelt et al (1999) cannot easily explain that syllables are constructed by rule in Dutch and at the same time account for their retrieval as chunks in Mandarin (see Cholin, Schiller & Levelt, 2004 for discussion). Likewise, the successor's of Dell's (1986) English grounded theory (e.g., Sevald, Dell & Cole, 1995) cannot retain the treatment of syllables as unselected structural units and also account for the status of Mandarin syllables as selected chunks. Both theories would need to specify different versions for each language. Yet another version would be needed to account for the Japanese findings of Kureta et al., (2006) which suggest that morae are the proximate units in that language. Although such language specific implementations are obviously useful, they do not meet the goal of providing a general theory. We propose instead that theories of language production need to be cast in terms of principles rather than stipulated roles for particular units. One such principle is the *proximate units principle* that we have introduced in this article, whereby the first selected phonological units, independently of their type, serve distinctive, primary functions in phonological encoding. It remains to be seen whether the proximate units approach holds up within groups of related languages, whether it generalizes successfully to less researched languages such as Japanese, and whether speakers of other languages are as firmly locked into particular modes as Mandarin speakers appear to be.

Acknowledgments

References

Alario F-X, Perre L, Castel C, Ziegler JC. The role of orthography in speech production revisited. Cognition 2007;102:464–475. [PubMed: 16545792]

Bock K. A sketchbook of production problems. Journal of Psycholinguistic Research Special Issue: Sentence processing 1991;20:141–160.

Chen J-Y. (A small corpus of speech errors in Mandarin Chinese and their classification). The World of Chinese Language 1993;69:26–41.

Chen J-Y. The representation and processing of tone in Mandarin Chinese: Evidence from slips of the tongue. Applied Psycholinguistics 1999;20:289–301.

Chen J-Y. Syllable errors from naturalistic slips of the tongue in Mandarin Chinese. Psychologia: An International Journal of Psychology in the Orient Special Issue: Cognitive processing of the Japanese and Chinese languages II 2000;43:15–26.

Chen J-Y, Chen T-M, Dell GS. Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. Journal of Memory & Language 2002;46:751–781.

Chen J-Y, Chen T-M. Form encoding in Chinese word production does not involve morphemes. Language and Cognitive Processes 2007;22:1001–1020.

Chen J-Y, Lin W-C, Ferrand L. Masked priming of the syllable in Mandarin Chinese syllable production. Chinese Journal of Psychology 2003;45:107–120.

Chen, J-Y.; O'Seaghdha, PG.; Liu, K-H. Phonological syllables control segmental encoding in Mandarin Chinese.. Poster presented at the Annual Meeting of the Psychonomic Society; Long Beach, CA. Nov. 2007

Chen, MY. Tone sandhi: Patterns across Chinese dialects. Cambridge University Press; 2000.

Chen T-M, Dell GS, Chen J-Y. A cross-linguistic study of phonological units: Syllables emerge from the statistics of Mandarin Chinese but not from the statistics of English. Proceedings of the Cognitive Science Society 2004:216–220.

Chen T-M, Chen J-Y. Morphological encoding in the production of compound words in Mandarin Chinese. Journal of Memory and Language 2006;54:491–514.

Chen, J-Y.; Chen, T-M.; O'Seaghdha, PG. Word form encoding in Chinese begins with the syllable: Further evidence from picture naming with masked primes.. Poster presented at the Annual Meeting of the Psychonomic Society; Boston, MA. Nov. 2009

Cheng, C-C. A synchronic phonology of Mandarin Chinese. Mouton; The Hague: 1973.

Cholin J, Schiller NO, Levelt WJM. The preparation of syllables in speech production. Journal of Memory and Language 2004;50:47–61.

Cholin J. The mental syllabary in speech production: An integration of different approaches and domains. Aphasiology 2008;22:1127–1141.

Costa A, Sebastian-Galles N. Abstract phonological structure in language production: Evidence from Spanish. Journal of Experimental Psychology: Learning, Memory and Cognition 1998;24:886–903.

Damian MF, Bowers JS. Effects of orthography on speech production in a form-preparation paradigm. Journal of Memory and Language 2003;49:119–132.

Dell GS. A spreading-activation theory of retrieval in sentence production. Psychological Review 1986;93:283–321. [PubMed: 3749399]

Dell, GS. Speaking and miss-speaking.. In: Gleitman, LR.; Liberman, M., editors. An Invitation to Cognitive Science. Vol 1: Language. MIT Press; Cambridge, MA: 1995. p. 183-208.

Dell GS, O'Seaghdha PG. Mediated and convergent lexical priming in language production: A comment on Levelt et al. (1991). Psychological Review 1991;98:604–614. [PubMed: 1961775]

Eickhoff SB, Heim S, Zilles K, Amunts K. A systems perspective on the effective connectivity of overt speech production. Philosophical Transactions of the Royal Society A 2009;367:2399–2421.

Ferrand L, Segui J, Grainger J. Masked priming of word and picture naming: The role of syllabic units. Journal of Memory and Language 1996;35:708–723.

Ferrand L, Segui J. The syllable's role in speech production: Are syllables chunks, schemas, or both? Psychonomic Bulletin & Review 1998;5:253–258.

Hickok G, Poeppel D. The cortical organization of speech processing. Nature Reviews Neuroscience 2007;8:393–402.

Jescheniak J, Schriefers H. Priming effects from phonologically related distractors in picture-word interference. Quarterly Journal of Experimental Psychology 2001;54A:371–382. [PubMed: 11394052]

Kubozono H. The mora and syllable structure in Japanese: Evidence from speech errors. Language and Speech 1989;32:249–278.

Kuo, F-L. Ph.D. thesis. University of Illinois at Urbana-Champaign; 1994. Aspects of segmental phonology and Chinese syllable structure..

Kureta Y, Fushimi T, Tatsumi IF. The functional unit of phonological encoding: Evidence for moraic representation in native Japanese speakers. Journal of Experimental Psychology: Learning, Memory and Cognition 2006;32:1102–1119.

Laganaro M, Alario F-X. On the locus of the syllable frequency effect in speech production. Journal of Memory and Language 2006;55:178–196.

Levelt, WJM. Speaking: From intention to articulation. MIT Press; Cambridge, MA: 1989.

Levelt WJM, Roelofs A, Meyer AS. A theory of lexical access in speech production. Behavioral & Brain Sciences 1999;22:1–75. [PubMed: 11301520]

Levelt WJM. Spoken word production: A theory of lexical access. Proceedings Of The National Academy Of Sciences 2001;98:13464–13471.

Meyer AS. The time course of phonological encoding in language production: The encoding of successive syllables of a word. Journal of Memory & Language 1990;29:524–545.

Meyer AS. The time course of phonological encoding in language production: Phonological encoding inside a syllable. Journal of Memory & Language 1991;30:69–89.

Meyer AS, Schriefers H. Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. Journal of Experimental Psychology: Learning, Memory and Cognition 1991;17:1146–1160.

Oppenheim GM, Dell GS. Inner speech slips exhibit lexical bias, but not the phonemic similarity effect. Cognition 2008;106:528–537. [PubMed: 17407776]

O'Seaghdha, PG.; Chen, J-Y. Toward a language-general account of word production: The proximate units principle.. In: Taatgen, NA.; van Rijn, H., editors. Proceedings of the 31st Annual Conference of the Cognitive Science Society. Cognitive Science Society; Austin, TX: 2009. p. 68-73.

O'Seaghdha PG, Marin JW. Phonological competition and cooperation in form-related priming: Sequential and nonsequential processes in word production. Journal of Experimental Psychology: Human Perception & Performance 2000;26:57–73. [PubMed: 10696605]

O'Seaghdha, PG.; Schuster, K.; Chen, J-Y. Integrated representation of novel spoken compounds.. Paper presented at the Psychonomic Society meeting; Vancouver, CANADA. 2003.

Otake T, Hatano G, Cutler A, Mehler J. Mora or syllable? Speech segmentation in Japanese. Journal of Memory and Language 1993;32:258–278.

Packard, JL. The morphology of Chinese: A linguistic and cognitive approach. Cambridge University Press; 2000.

Perret C, Bonin P, Meot A. Syllabic priming effects in picture naming in French: Lost in the sea! Experimental Psychology 2006;53:95–104. [PubMed: 16909933]

Rapp B, Goldrick M. Discreteness and interactivity in spoken word production. Psychological Review 2000;107:460–499. [PubMed: 10941277]

Roelofs A. The WEAVER model of word form encoding in speech production. Cognition 1997;64:249–284. [PubMed: 9426503]

Roelofs A, Meyer AS. Metrical structure in planning the production of spoken words. Journal of Experimental Psychology: Leanring, Memory and Cognition 1998;24:922–939.

Roelofs A. Phonological segments and features as planning units in speech production. Language & Cognitive Processes 1999;14:173–200.

Roelofs A. The influence of spelling on phonological encoding in word reading, object naming, and word generation. Psychonomic Bulletin & Review 2006;13:33–37. [PubMed: 16724765]

Santiago J. Implicit priming of picture naming: A theoretical and methodological note on the implicit priming task. Psicologica 2000;21:39–59.

Schiller NO. The effect of visually masked syllable primes on the naming latencies of words and pictures. Journal of Memory & Language 1998;39:484–507.

Schiller, NO.; Costa, A.; Colomé, A. Phonological encoding of single words: In search of the lost syllable.. In: Gussenhoven, C.; Warner, N., editors. Papers in Laboratory Phonology VII. Mouton de Gruyter; Berlin: 2002.

Shattuck-Hufnagel, S. Sublexical units and suprasegmental structure in speech production planning.. In: MacNeilage, PF., editor. The production of speech. Springer-Verlag; New York: 1983. p. 109-136.

Sevald CA, Dell GS. The sequential cuing effect in speech production. Cognition 1994;53:91–127. [PubMed: 7805353]

Sevald CA, Dell GS, Cole J. Syllable structure in speech production: Are syllables chunks or schemas? Journal of Memory & Language 1995;34:807–820.

Wong AW-K, Chen H-C. Processing segmental and prosodic information in Cantonese word production. Journal of Experimental Psychology: Leanring, Memory and Cognition 2008;34:1172–1190.

(A) Production of a Mandarin Monosyllable   **(B) Production of an English Monosyllable**
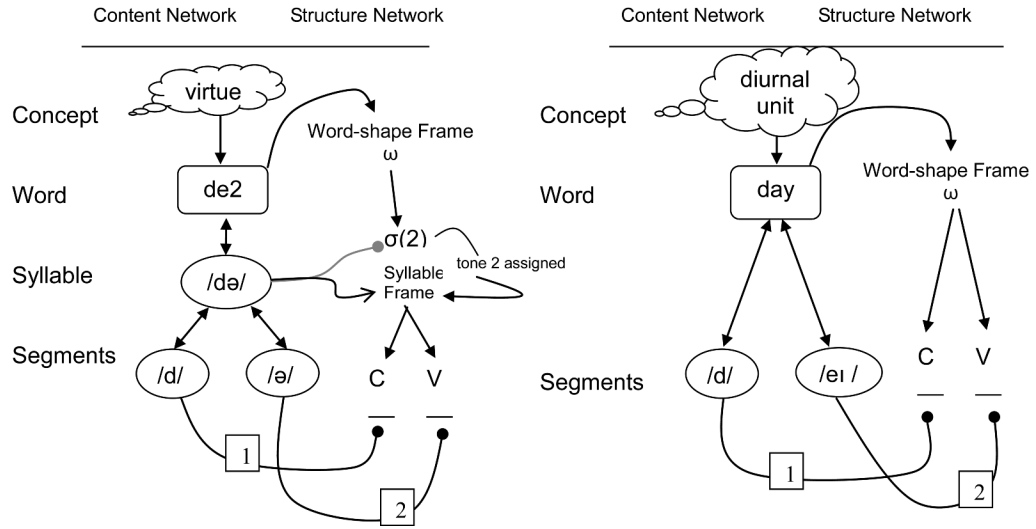
Figure 1.
Production of a CV monosyllable in a) Mandarin and b) English. In Mandarin but not in English syllables are retrieved as chunks. Arrows signify activation. Button terminals signify assignment of contents to structures. See text for additional details.

Table 1

Experiment 1: Mean response times in ms. (M) in Homogeneous and Heterogeneous contexts, with error rates (E%), standard errors (SE), and preparation effects (Difference).

Set		Homogeneous	Heterogeneous	Difference
/m/	M	597	587	-10
	E%	.04	.04	
	SE	7.98	7.46	
/d/	M	601	601	0
	E%	.03	.03	
	SE	8.06	9.16	
/sh/	M	630	642	12
	E%	.02	.02	
	SE	8.16	7.87	
/l/	M	588	598	10
	E%	.02	.03	
	SE	9.48	8.06	
Overall	M	604	607	3
	E%	.03	.03	
	SE	4.38	4.39	

Table 2

Experiments 2-4: Mean response times to Mandarin monosyllables in ms. (M) in Homo(geneous) and Hetero (geneous) contexts, with error rates (E%), standard errors (SE), and preparation effects (Difference).

Set		**EXPT. 2: Compounds**			**EXPT. 3: Monosyllables**			**EXPT. 4: Symbol Prompts**			
		Homo-	Hetero-	Difference	Homo-	Hetero-	Difference	Homo-	Hetero-	Difference	Overall
/m/	M	536	518	-18	560	560	0	617	597	-20	-13
	E%	.01	.01		.02	.01		.08	.03		
	SE	8.70	7.96		8.60	7.51		10.64	10.30		
/d/	M	550	550	0	601	607	6	589	599	10	5
	E%	.04	.03		.05	.02		.02	.02		
	SE	9.69	8.18		7.51	7.86		11.08	10.38		
/sh/	M	596	606	10	664	661	-3	625	637	12	6
	E%	.02	.00		.03	.03		.03	.02		
	SE	8.54	8.72		8.73	6.92		11.33	8.02		
/l/	M	533	517	-17	560	566	6	587	602	15	1
	E%	.01	.01		.01	.03		.02	.03		
	SE	9.23	7.92		7.83	7.76		10.11	12.36		
Overall	M	554	548	-6	596	599	3	605	609	4	0
	E%	.02	.01		.03	.03		.04	.03		
	SE	4.94	5.06		8.39	5.03		5.52	5.31		

Table 3

Experiments 5 and 6: Mean response times to English monosyllables in ms. (M) in Homogeneous and Heterogeneous contexts, with error rates (E%), standard errors (SE), and preparation effects (Difference)

Set		Expt 5: Word Prompts			Expt 6: Picture Prompts		
		Homogeneous	Heterogeneous	Difference	Homogeneous	Heterogeneous	Difference
/p/	M	680	693	13	564	565	1
	E%	.07	.05		.07	.05	
	SE	14.77	14.88		14.56	14.51	
/d/	M	643	670	27	553	577	24
	E%	.05	.06		.06	.05	
	SE	18.26	16.84		15.53	12.37	
/s/	M	658	666	9	542	557	15
	E%	.09	.06		.06	.05	
	SE	22.86	22.59		15.75	13.36	
/r/	M	685	686	1	576	600	25
	E%	.10	.07		.06	.07	
	SE	17.68	17.11		15.59	13.11	
Overall	M	666	679	12	559	575	16
	E%	.08	.06		.06	.05	
	SE	17.25	16.77		12.50	13.98	

Table 4

Experiment 7: Results by Group, by Set, and by Block: Mean response times in ms. (M) for Syllable and Onset Conditions in Homogeneous and Heterogeneous Contexts, with error rates (E%), standard errors (SE), and preparation effects (Diff = Difference).

| Group | | Syllable | | | Onset | | |
		Homogeneous	Heterogeneous	Diff	Homogeneous	Heterogeneous	Diff
1	M	588	596	8	582	584	2
	E%	.07	.06		.04	.03	
	SE	4.55	4.55		4.76	4.08	
2	M	578	590	12	595	594	-1
	E%	.05	.06		.06	.06	
	SE	5.27	4.23		5.08	4.24	
3	M	594	602	7	616	621	5
	E%	.06	.07		.08	.06	
	SE	5.25	4.74		5.22	4.59	
Set							
/f/	M	576	590	14	597	598	1
	E%	.06	.09		.09	.07	
	SE	4.99	4.39		4.80	4.68	

/k/	M	587	591	4	590	590	0
	E%	.06	.05		.04	.04	
	SE	5.06	4.54		5.32	4.23	
/p/	M	597	607	10	606	611	5
	E%	.06	.05		.05	.05	
	SE	4.98	4.48		5.37	4.54	
Block							
1	M	604	607	3	609	611	2
	E%	.08	.08		.07	.08	
	SE	4.83	4.19		5.21	4.47	
2	M	585	594	9	596	599	3
	E%	.06	.07		.06	.06	
	SE	5.00	4.70		4.99	4.45	
3	M	572	587	15	588	588	0
	E%	.06	.06		.05	.04	
	SE	4.92	4.48		5.23	4.51	
Blocks 2 & 3	M	578	590	12	592	594	2
	E%	.06	.06		.06	.05	
	SE	3.53	3.25		3.61	3.18	
All	M	587	596	9	597	599	2
	E%	.07	.07		.06	.06	
	SE	2.92	2.61		3.00	2.63	

Table 5

English Homogeneous Response Set /d/ and Picture Prompts

Picture	Response
	Day
	Dew
	Dough
	Dye

IV	/sh/	
法庭 - 律師 fa3-ting2 - lv4-shi1 'court – lawyer'	科學 - 實驗 ke1-xue2 - shi2-yan4 'science – experiment'	
整形 - 拉皮 zheng3-xing2 - la1-pi2 'plastic surgery – face lift'	fan4-xia4 - she3-qi4 'put down – give up'	
分開 - 離別 fen1-kai1 - li2-bie2 'separate – leave'	緊急 - 煞車 jin3-ji2 - sha4-che1 'emergency – brake'	
小吃 - 滷肉 xiao3-chi1 - lu3-lou4	毛筆 - 書法 mao2-bi3 - shu1-fa3	

/l/	/sh/	
'street food – marinated meat'	'brush – calligraphy'	

/V

法←→律
fa3←→lv4
'law'

拖→拉
tuo1→la1
'pull'

N

qv4←li2
'leave'

rou4←lu3
'marinated meat'

IV

嚴 – 律
yan2 - lv4
'strict-law'

提 – 拉
ti2 - la1
'lift-pull'

和

IV

近 – 離
jin4 - li2
'near-leave'

yao4 - lu3
'medicine-stew'

N

la1
'pull'

/V

@ - 離

li2
'leave'

/l/

湖
#一
lu3
'stew'

5

Monosyllable Pairs in English Experiment 5

		Homogeneous			
	Set	/d/	/p/	/r/	/s/
Heterogeneous	1	night-day	snap-pea	grain-rye	pig-sow
	2	tint-dye	church-pew	light-ray	boat-sea
	3	bread-dough	cash-pay	sauce-roux	gasp-sigh
	4	wet-dew	fruit-pie	line-row	talk-say

p1

比方 - 譬喻
bi3fang1-pi4yu4
example - metaphor

運動 - 跑步
yun4dong4-pao3bu4
exercise - jogging

p1

蚊蟲-撲滅
wen2chong2-pu1mie4
mosquito - exterminate

p2		
	xiao1ye4-pao4mian4 night snack - instant noodles	
	流行 - 普遍 liu2xing2-pu3pian4 popular - general	

p3

流氓-痞子
liu2mang2-pi3zi5
rogue - ruffian

故障-抛锚
gu4zhang4-pao1mao2
malfunction - break down

峡谷-瀑布
xia2gu3-pu4bu4
gorge - waterfall