Brain Research 1648 (2016) 101–109

Contents lists available at ScienceDirect

Brain Research

journal homepage: www.elsevier.com/locate/brainres

Research report

Syllabic encoding during overt speech production in Cantonese: Evidence from temporal brain responses

Andus Wing-Kuen Wong [a,1], Jie Wang [b,1], Tin-Yan Ng [b], Hsuan-Chih Chen [b,*]

[a] Nam Shan Psychology Laboratory, Department of Applied Social Sciences, City University of Hong Kong, Nam Shan Chuen Road, Kowloon, Hong Kong, China
[b] Department of Psychology, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, China

ARTICLE INFO

Article history:
Received 21 December 2015
Received in revised form
10 June 2016
Accepted 19 July 2016
Available online 20 July 2016

Keywords:
Phonological encoding
Cantonese speech production
Event-related brain potentials (ERP)

ABSTRACT

The time course of phonological encoding in overt Cantonese disyllabic word production was investigated using a picture-word interference task with concurrent recording of the event-related brain potentials (ERPs). Participants were asked to name aloud individually presented pictures and ignore a distracting Chinese character. Participants' naming responses were faster, relative to an unrelated control, when the distractor overlapped with the target's word-initial or word-final syllables. Furthermore, ERP waves in the syllable-related conditions were more positive-going than those in the unrelated control conditions from 500 ms to 650 ms post target onset (i.e., a late positivity). The mean and peak amplitudes of this late positivity correlated with the size of phonological facilitation. More importantly, the onset of the late positivity associated with word-initial syllable priming was 44 ms earlier than that associated with word-final syllable priming, suggesting that phonological encoding in overt speech runs incrementally and the encoding duration for one syllable unit is approximately 44 ms. Although the size of effective phonological units might vary across languages, as suggested by previous speech production studies, the present data indicate that the incremental nature of phonological encoding is a universal mechanism.

1. Introduction

Scholars have long been interested in the commonalities, as well as differences, across different natural languages. Findings from the relevant work have provided important insights for our understanding of the relationship between language processing and other human cognitive functioning. In the past few decades, various approaches have been adopted to investigate the language-universal and language-specific aspects in the processing of human language. One key topic in this line of research relates to the cognitive mechanisms underlying speech production. Specifically, to what extent that the processes of retrieving and formulating the phonological content of a spoken word are similar across distinctive languages has recently been a topic of interest.

To produce a spoken word, the phonological form of the target needs to be prepared to guide subsequent articulatory movements (Caramazza, 1997; Dell, 1986, 1988; Levelt, 1989; Levelt et al., 1999). The processes of phonological form preparation have been referred to as phonological encoding. According to one prominent theory of speech production, the WEAVER model (Roelofs, 1992, 1997), phonological encoding starts after the target's lexical semantic-syntactic properties have been determined, and involves specifying the target's phonological segments (phonemes), number of syllables, and stress location. The output of phonological encoding is a syllabified phonological representation with its constituent segments properly ordered. Furthermore, the model assumes that the encoding process operates incrementally from the beginning to the end of the target. Accordingly, phonological encoding of the word-initial segments is initiated prior to the encoding of the word-final segments, and this assumption holds for both monosyllabic and multisyllabic target words.

One of the earliest experimental evidences supporting the incremental view of phonological encoding was reported by Meyer (1990) using the implicit priming paradigm. In a word-associative naming task, knowing the first syllable of the targets facilitated Dutch speakers' naming response, while knowing the second syllable had no such an effect. The failure in preparing non-initial syllables was attributed to the rightward direction of phonological

Abbreviations: ANOVA, analysis of variance; EEG, electroencephalogram; EOG, electrooculogram; ERP, event-related brain potential; PWI, picture-word interference; ROI, region of interest; RT, reaction time; WEAVER, word-form encoding by activation and verification

* Corresponding author.
E-mail addresses: andus.wong@cityu.edu.hk (A.-K. Wong), jwang@psy.cuhk.edu.hk (J. Wang), htyng@psy.cuhk.edu.hk (T.-Y. Ng), hcchen@psy.cuhk.edu.hk (H.-C. Chen).

[1] Author Contributions: AWKW and JW contributed equally to this work and should be considered as co-first authors.

http://dx.doi.org/10.1016/j.brainres.2016.07.032
0006-8993/

encoding. Meyer (1991) further tested this implicit priming effect at the segment level, and found a similar result (i.e., a significant onset effect but no rhyme effect). In addition, consistent results (Meyer and Schriefers, 1991) were found in a picture-word interference (PWI) task, in which participants were asked to name aloud pictures and ignore auditory word distractors. When the target picture name and the distractor shared similar word-initial segments, facilitation in naming latency, relative to an unrelated control, was obtained in various stimulus onset asynchronies (SOAs) as early as when SOA was −150 ms. In contrast, word-end related distractors produced facilitation effects only when the shared segments were presented at or shortly after the onset of picture presentation (SOA=0 or +150 ms), suggesting that phonological encoding of the word-initial segments is initiated prior to the encoding of the word-final segments.

To examine the time course of phonological encoding, Wheeldon and Levelt (1995) asked participants to monitor their internal speech for pre-designated target segments. Target segments were monitored significantly faster when they appeared in the onset position of the first syllable of the target-bearing word than in that of the second syllable, and the difference in monitoring latency between the two varied from 37 ms to 150 ms across experimental conditions. Using event-related brain potentials (ERPs) to signal motor preparation in real time during the speech monitoring task, van Turennout et al. (1997) reported that the initial segment of a target word was encoded 80 ms earlier than its final segment (1.5 syllables per word on average). Similarly, in a stress monitoring task with disyllabic targets, Schiller (2006) found a 58 ms difference in ERP peak latencies between the word-initial stress condition and the word-final stress condition. These findings suggest that the time lag between the encoding of the first and the second syllable was approximately 50 ms.

Nevertheless, it is not self-evident that the above findings from European languages can easily be generalized to languages with distinctively different phonological features (e.g., Chinese). Indeed, studies on Mandarin and Cantonese spoken word production failed to find the evidence supporting incremental phonological encoding at the segment level. While no onset effect on naming performance was found (Mandarin: Chen et al., 2002; Chen and Chen, 2013; O'Seaghdha et al., 2010; Zhang, 2008; Zhang and Yang, 2005; Cantonese: Wong and Chen, 2008; Wong et al., 2012; see Li et al., 2015; Verdonschot et al., 2013 for exceptions), manipulations on the rhyme position either facilitated or inhibited naming responses in different studies (Mandarin: Zhang, 2008; Li et al., 2015; Cantonese: Wong and Chen, 2008, 2009, 2015), which cannot be explained by the assumption of incremental phonological encoding at the segment level. Given that null effects of the onset consonant have been reported in Mandarin (Chen et al., 2002) and Cantonese Chinese (Wong and Chen, 2008; Wong et al., 2012), and that significant effects were obtained with the onset consonant in European languages such as Dutch and English (Meyer, 1990), it has been argued that the size of phonological unit is different across languages with distinctive phonological systems (O'Seaghdha, 2015; O'Seaghdha et al., 2010; Roelofs, 2015). However, whether the assumption of incremental phonological encoding is language-universal which also persists at the syllable level of Chinese spoken word production has not been systematically studied. Therefore, the present study was conducted to investigate the time course of phonological encoding of the first and the second syllable in Cantonese disyllabic word production.

More critically, it is not clear to what extent the findings from the speech monitoring paradigm (with or without concurrent ERP recording) reflect the time course of normal speech production processes, since an overt speech response is not required in the monitoring task. Previous ERP studies typically adopted covert speech production instead of overt speech production to avoid motor artifacts induced by oral-facial movements. However, most of the phonological encoding processes of interest should have been completed 150 ms prior to the onset of articulation (Indefrey and Levelt, 2004). Hence, it is plausible in principle to obtain artifact-free brain responses in the early time windows. Fortunately, recent ERP studies have shown that it is possible to combine an overt speech task with concurrent ERPs recording to investigate the time course of various speech production processes (Blackford et al., 2012; Costa et al., 2009; Dell'Acqua et al., 2010; Ganushchak et al., 2011; Qu et al., 2012; Yu et al., 2014). Zhu et al. (2015), for instance, investigated the time course of semantic and phonological processes in overt Mandarin speech production using a PWI task together with concurrent ERP recording. A significant semantic-related ERP effect was found in the 250–450 ms time window after picture onset, whereas a significant phonological-related ERP effect was obtained in the 450–600 ms time window. Of particular relevance to the present study is that the phonological related distractors used in Zhu et al. (2015) shared a similar word-initial atonal syllable (i.e., a syllable unit without the tone specified) with the target, indicating that the encoding of the word-initial segments/atonal syllable in Mandarin speech occurs in the time window between 450 and 600 ms post target onset. However, the phonological related distractors used in Zhu et al. (2015) always overlapped with the targets in their word-initial syllable and no manipulation was implemented to prime other syllable positions, so the results of their study do not allow one to examine the assumption of incremental phonological encoding and its precise time course.

There is yet no electrophysiological evidence available regarding the detailed time course of phonological encoding during overt speech production. One relevant study was conducted by Yu et al. (2014), where participants were asked to name aloud individually presented pictures in Mandarin with concurrent ERP recording. Significant modulations in ERP signals were found when the target picture was preceded by a prime picture where their names shared an identical segment in the word-initial syllable position (e.g., /ling3 dai4/meaning “tie” and/la4 zhu2/meaning “candle”; Experiment 1) or in the word-final syllable position (e.g., /chou1 ti4/meaning “drawer” and/jiao4 tang2/meaning “church building”; Experiment 2) (Yu et al., 2014). Notably, the modulation effects in these two priming conditions were both observed within the time window of 180 ms to 300 ms post target presentation. However, no further details were reported regarding the time course of the two modulation effects (e.g., their onset or peak latencies). Furthermore, the modulation effects were obtained across two experiments using different participants and materials; a direct comparison between the two effects was therefore difficult to be made. So it remains unclear of whether the results of Yu et al. (2014) are consistent with the serial, or a parallel, view of phonological encoding. Therefore, the present study was conducted to directly investigate the detailed time course of phonological encoding in spoken word production using the PWI task, a frequently used paradigm in the literature of language production, with concurrent ERP measurement.

In the context of a PWI task, the presence of a distractor would generally slow down participant's naming response relative to a condition where there is no distractor (Schriefers et al., 1990). One typical finding is that participants' naming responses would be faster when the target and distractor shared similar phonological contents than if they are phonologically unrelated (Meyer and Schriefers, 1991), and such phonological facilitation effect has been argued to be arising from the phonological encoding stage (e.g., Damian and Martin (1999)). In the present study, the PWI paradigm was adopted with simultaneous EEG recording during overt speech to examine the time lag between the encoding of the first and the second syllable in Cantonese disyllabic word production.

Please cite this article as: Wong, A.W.-K., et al., Syllabic encoding during overt speech production in Cantonese: Evidence from temporal brain responses. Brain Research (2016), http://dx.doi.org/10.1016/j.brainres.2016.07.032