Neuropsychologia 64 (2014) 71–84

Contents lists available at ScienceDirect

Neuropsychologia

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

How and when predictability interacts with accentuation in temporally selective attention during speech comprehension

Xiaoqing Li [a,*], Yong Lu [b], Haiyan Zhao [a]

[a] Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China
[b] Academy of Psychology and Behavior, Tianjin Normal University, Tianjin, China

ARTICLE INFO

Article history:
Received 4 March 2014
Received in revised form
10 September 2014
Accepted 12 September 2014
Available online 22 September 2014

Keywords:
Speech processing
Temporally selective attention
Accentuation
Predictability

ABSTRACT

The present study used EEG to investigate how and when top-down prediction interacts with bottom-up acoustic signals in temporally selective attention during speech comprehension. Mandarin Chinese spoken sentences were used as stimuli. We systematically manipulated the predictability and de/accentuation of the critical words in the sentence context. Meanwhile, a linguistic attention probe 'ba' was presented concurrently with the critical words or not. The results showed that, first, words with a linguistic attention probe elicited a larger N1 than those without a probe. The latency of this N1 effect was shortened for accented or lowly predictable words, indicating more attentional resources allocated to these words. Importantly, prediction and accentuation showed a complementary interplay on the latency of this N1 effect, demonstrating that when the words had already attracted attention due to low predictability or due to the presence of pitch accent, the other factor did not modulate attention allocation anymore. Second, relative to the lowly predictable words, the highly predictable words elicited a reduced N400 and enhanced gamma-band power increases, especially under the accented conditions; moreover, under the accented conditions, shorter N1 peak-latency was found to correlate with larger gamma-band power enhancement, which indicates that a close relationship might exist between early selective attention and later semantic integration. Finally, the interaction between top-down selective attention (driven by prediction) and bottom-up selective attention (driven by accentuation) occurred before lexical-semantic processing, namely before the N400 effect evoked by predictability, which was discussed with regard to the language comprehension models.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Speech signal provides a large amount of information and unfolds rapidly in time, which presents significant challenges to the auditory perception and comprehension systems. Listeners must determine which time point of the rapidly changing acoustic signals needs to be processed in detail. Temporally selective attention, therefore, would be critical for speech comprehension, since it can help to allocate attentional resources to time windows that contain more important information (Astheimer and Sanders, 2009). A striking feature of selective attention is that it depends not only on "bottom-up" signals, such as sensory salience, but also on "top-down" information, such as our prior knowledge or predictions. With regard to the relationship between these top-down and bottom-up processes, previous studies mainly focus on visual–spatial attention. How people use temporally selective attention to process speech is still a relatively underdeveloped area. Thus, the present study aimed to investigate how the bottom-up process driven by acoustic salience (namely, accentuation) and the top-down process driven by prior prediction interact with each other in temporally selective attention during speech comprehension.

Accentuation is one type of prosodic information in the speech signal, which reflects the relative prominence of a particular syllable, word, or phrase in an utterance realized mainly by modulations of pitch (Shattuck-Hufnagel and Turk, 1996).

In West-Germanic languages, it is by now well-known that speakers tend to place a pitch accent on new or focused information, while leaving given information de-accented (Gussenhoven, 1983; Ladd, 1996). In Chinese, the new or focused information is also encoded via accentuation which is mainly realized by the expansion of the pitch range of lexical tone (Chen, 2006; Chen and Gussenhoven, 2008; Xu, 1999; Jia et al., 2008, 2006; Liu and Xu, 2006; Wang et al., 2002). Previous psycholinguistic studies on accentuation mainly focused on the correspondence between accentuation and information structure. Behavioral studies found that speech processing is facilitated when new information is accented and given information de-accented (e.g. Bock and

* Correspondence to: Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, No. 16 Lincui Road, Chaoyang District, Beijing 100101, China. Tel.: +86 10 64864012; fax: +86 10 64872070.
E-mail address: lixq@psych.ac.cn (X. Li).

http://dx.doi.org/10.1016/j.neuropsychologia.2014.09.020
0028-3932/© 2014 Elsevier Ltd. All rights reserved.

Mazzella, 1983; Dahan et al., 2002; Terken and Noteboom, 1987). The studies using EEG (electroencephalogram) also reported immediate ERP (event related potential) effects (broadly-distributed negativity, N400, or P300) for missing pitch accent on new information or superfluous pitch accent on given information (e.g., Dimitrova et al., 2012; Hruska et al., 2000; Johnson et al., 2003, Li et al., 2008b; Magne et al., 2005; Toepel et al., 2007). Those results suggest that accentuation can influence the ease by which the current speech signal is processed.

As to the specific mechanisms by which accentuation affects speech processing, some researchers propose that accentuation can modulate attention allocation (e.g., Cutler, 1976; Li and Ren, 2012; Sanford et al., 2006). For example, using phoneme monitoring task, Cutler (1976) showed heightened attention (as indicated by faster phoneme monitoring responses) to a word that received a pitch accent. Subsequently, using change detection task, Sanford et al. (2006) also found that the ability of listeners to detect a one-word alteration between the twice-presented spoken discourses was superior under the contrastive accent condition than that under the non-contrastive accent condition. Using EEG, Li and Ren (2012) also found that the semantically incongruent words elicited a larger N400 than the semantically congruent words when the corresponding words were accented; however, no significant difference was observed between the incongruent and congruent words when they were de-accented. Those results indicate that accentuation can modulate listeners' selective attention process. Accentuation guides the listeners to allocate more attention to accented information and take deeper processing. Therefore, listeners are more able to detect the presence of semantic incongruence, word alteration, or phoneme probe under the accented condition (Cutler, 1976; Li and Ren, 2012; Sanford et al., 2006).

Besides the bottom-up sensory signal such as accentuation, prior knowledge or predictions also influence the process of language comprehension. The effect of predictability on language comprehension has been studied extensively in both spoken language comprehension and reading. Eye tracking studies revealed that highly predictable words are read more quickly and skipped more often than lowly predictable words (e.g., Frisson et al., 2005). The ERP studies also demonstrated that, in a sentence or discourse context, the word with a high-level of cloze-probability (namely, a high-level of predictability) elicits a reduced N400 compared with the word with a low-level of cloze-probability (e.g., DeLong et al., 2005; Van Berkum et al., 2005; Federmeier, 2007; Laszlo and Federmeier, 2009; Thornhill and Van Petten, 2012). Those results indicate that the highly predictable words are more easily processed and integrated into the sentence or discourse context.

Some fMRI or MEG studies also examined how predictability influences early perceptual processing and found that predicted stimuli evoke reduced neural responses in the early visual/auditory cortex (Alink et al., 2010; den Ouden et al., 2010; Todorovic et al., 2011; Sohoglu et al., 2012). For example, Sohoglu et al., 2012 manipulated predictability of speech content by presenting matching, mismatching, or neutral text before speech onset. They found that the provision of prior knowledge reduces activity in the superior temporal gyrus that has been considered to be involved in perceptual aspects of speech processing. Although the precise relationship between predictive sensory coding and attention is still the subject of ongoing debate, the sensory attenuation of predicted signals is consistent with the possibility that processors might direct less attention to predicted external signals. However, other studies also provided inconsistent results by revealing that prediction sometimes seems to enhance rather than reduce sensory signals (Doherty et al., 2005; Chaumon et al., 2008). To resolve the inconsistency in the above results, Kok et al. (2012) further examined how predictability and cued spatial attention affect early visual perceptual processing. They found that, at the unattended location, predicted stimuli reduce neural response in the early visual cortex; in contrast, at the attended location, predicted stimuli enhance neural response in the early visual cortex. That is, the effect of predictability on early perceptual processing is modulated by the amount of attentional resources already allocated to the external stimuli.

Until now, only a few studies directly examined how predictability influences temporally selective attention during speech processing. In an ERP study, Astheimer and Sanders used an attention probe paradigm to investigate whether listeners allocate attentional resources to the time windows that contain highly important acoustic information, such as word onset. Attention probes were presented concurrently with word onsets, beginning 50 and 100 ms before and after word onsets, and at random control intervals. They found that linguistic probe 'ba' presented at the word onset elicited larger amplitude N1 than probes presented at other time points, suggesting that listeners direct attention to moments that contain word onsets (Astheimer and Sanders, 2009). Subsequently, Astheimer and Sanders (2011) further explored the reason that listeners attend to word onsets in speech. Based on transitional probabilities, word onsets are relatively unpredictable (Aslin et al., 1999). It might be that listeners tend to allocate more resources to times at which unpredictable information is presented, since unpredictable segments are highly informative. To test this hypothesis, they measured ERPs elicited by syllable onsets in an artificial language. The participants were required to listen to stream of artificial nonsense words arranged in pairs, such that the second word in each pair was completely predictable. After recognition training, the unpredictable first words elicited a larger N1. This enhancement was absent for the completely predictable second word in each pair. These results provided solid evidences for the fact that listeners are most likely to attend the segments in speech that are less predictable (Astheimer and Sanders, 2011).

Taken together, the above findings make it clear that both predictability and sensory salience, such as accentuation, influence temporally selective attention during speech processing. However, there are still questions needing to be explored further. First, although previous studies proved that accentuation plays a role in modulating attention allocation during spoken language comprehension, their findings are based on behavioral measures or on semantic congruence effect (e.g., Cutler, 1976; Li and Ren, 2012; Sanford et al., 2006). Therefore, these studies could not tell us whether accentuation can modulate selective attention at the early stage of information processing, such as before lexical-semantic processing or decision making. Second, with the help of artificial word training paradigm, the study conducted by Astheimer and Sanders (2011) demonstrated that listeners allocate more resources to the less predictable moment in the speech signal. However, we do not know whether, during natural speech comprehension, predictions derived from sentence context have the same effect on selective attention. Third and most importantly, it is completely unknown how, or even if, the bottom-up process driven by sensory salience and the top-down process driven by prior prediction interacts with each other in temporally selective attention during speech comprehension. If they do, at what functional stage does the top-down process begin to affect the bottom-up process?

As to the functional stage at which top-down knowledge interacts with bottom-up sensory signals, different models have been put forward. One proposal assumes that language processing is strictly feedforward, with semantic contextual information and bottom-up sensory information integrated only at a later decision stage (Fodor, 1983; McQueen et al., 2006; Norris et al., 2000). In contrast, the TRACE (McClelland and Elman, 1986; McClelland

دریافت فوری ← متن کامل مقاله

ISIArticles
مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات