How does dialectal experience modulate anticipatory speech processing?

Li, X.; Ren, G.; Zheng, Y.; Chen, Y.

How does dialectal experience modulate anticipatory speech processing?

Xiaoqing Li[a,b,*], Guiqin Ren[c], Yuanyi Zheng[a,b], Yiya Chen[d,*]

[a] CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China
[b] Department of Psychology, University of Chinese Academy of Sciences, Beijing 100101, China
[c] College of Psychology, Liaoning Normal University, Dalian 116029, China
[d] Leiden University Center for Linguistics (LUCL) & Leiden Institute for Brain and Cognition (LIBC), Leiden 2300 RA, The Netherlands

ABSTRACT

Context-based predictions facilitate speech processing. However, details of predictive processing mechanisms and how factors like language experience shape facilitative processing remain debated. This electroencephalograph study aimed to shed light on these issues by investigating the effect of dialectal experience on lexical prediction. Stimulus sentences were produced in three Mandarin Chinese dialects (home dialect, familiar regional dialect, and unfamiliar regional dialect). Critical nouns varied between strong and weak predictability. Only when listening to the home dialect was an enhanced ERP deflection observed at the transitive verb before a strongly predictable object noun (compared to the weakly predictable condition); the ERP amplitude at the verb correlated significantly with the following noun's predictability. The predictable object nouns elicited a reduced N400 in all three dialects but only an additional reduction in N1 in the home dialect condition. Conjointly, our results suggest that effortful meaning computation may partially underlie anticipatory lexical processing and that language experience modulates the way lexical prediction facilitates the early stage of acoustic/phonological processing (in addition to the later stage of meaning integration).

Introduction

It has been argued that the human brain, while performing cognitive activities, continuously anticipates incoming inputs with top-down predictions based on available information (Friston, 2005, 2010). Predictive processing provides a mechanism by which higher-level information is constantly transmitted to lower sensory levels so that top-down conceptual constraints are enabled to influence early sensory or perceptual stages of processing. It has also been posited that during language comprehension, readers/listeners use available cues incrementally and predict upcoming words, which is known as lexical prediction. This facilitates the processing of newly available bottom-up information at multiple levels of representation such as semantic, morpho-syntactic, phonological, and orthographic levels (Federmeier, 2007; Pickering & Garrod, 2013; Kuperberg & Jaeger, 2016; Levy, 2008). Lexical prediction is believed to be beneficial for speech comprehension, as everyday speech is rich in segmental/suprasegmental acoustic variations and is often filled with auditory ambient noise, which tends to impede the early stage of speech perception.

There is currently ample evidence for the facilitation effect of prediction on the semantic integration of a word into its preceding context

(see, e.g., Federmeier, 2007 for review). However, evidence for this facilitating effect on early pre-semantic (e.g., sensory or perceptual) stages of processing is inconsistent and less robust (see Nieuwland, 2019 for review). Moreover, although a range of factors (such as cognitive abilities and long-term knowledge) have been found to modulate context-based anticipatory language processing (Brothers, Swaab, & Traxler, 2017; Kroll & De Groot, 2005), the detailed mechanisms in how different factors enable/inhibit predictive language processing need further clarification.

One particularly interesting aspect of spoken language communication is that in everyday life, we not only use our native language but also are likely to encounter familiar/unfamiliar regional varieties that we understand but do not necessarily speak. Intuitively, cross-dialectal speech comprehension should be partly contingent on listeners' cognitive capacity to deal with acoustic variability in pronunciation due to dialectal variations. Predictive processing in such contexts may be even more critical for efficient speech communication but may also become less efficient due to the increased cognitive cost of sensory-perceptual processing and transmitting information between higher semantic/conceptual and lower sensory/perceptual levels of representations. Tapping into the role of dialectal experience in

predictive speech processing thus provides an unexplored frontier that promises to shed new light on our understanding of lexical prediction.

Lexical prediction in native language

Results of behavioral studies have established that lexical predictability facilitates language processing; predictable words are recognized faster than are unpredictable ones (e.g. Kamide, Altmann, & Haywood, 2003; Nakamura, Arai, & Mazuka, 2012). An increasing number of EEG studies add further evidence that words with strong predictability elicit an amplitude reduction in the N400 component compared to those that are only weakly predictable (e.g., DeLong, Urbach, & Kutas, 2005; Laszlo & Federmeier, 2009; Lau, Almeida, Hines, & Poeppel, 2009; Thornhill & Van Petten, 2012; see Federmeier, 2007 for review). Given that the N400 is a negative deflection commonly associated with meaning processing (Kutas & Hillyard, 1980), such an event-related potential (ERP) effect has been taken to reflect that word predictability can facilitate the semantic processing of an incoming new word due to *pre-activation* (i.e., prediction or early access of the upcoming word) (e.g., Kutas & Federmeier, 2000), ease of *integration* (i.e., unification of the incoming word meaning with the existing discourse context) (Hagoort, Hald, Bastiaansen, & Petersson, 2004), or the cascaded effects of both *access* and *integration* processes (Baggio & Hagoort, 2011; Nieuwland et al., 2019).

While the N400 amplitude reduction with highly predictable words suggests anticipatory lexical processing, it has been well-recognized that stronger evidence for anticipatory processing of target words comes from processing linguistic elements prior to the onset of these target words due to their pre-activation. DeLong et al. (2005) conducted an influential study as the first to present evidence of pre-activation by manipulating the congruency between the target nouns' phonological forms (e.g., *kite* vs. *airplane*) and their preceding indefinite articles (*a* vs. *an*). Their results showed a reduced ERP amplitude in the N400 time window as a function of high cloze probability not only for nouns (e.g., *kite*) but also for articles (e.g., *a*), lending support to the pre-activation of target nouns. Further evidence for lexical pre-activation came from studies that manipulated the grammatical gender markers of predicted target nouns and found evidence of mismatched gender markers before the following target nouns became available (e.g., Otten & Van Berkum, 2009; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005 for Dutch; Wicha et al., 2003, 2004 for Spanish). The replication attempts, however, failed to reproduce the same effect reported in DeLong et al.'s study (i.e., Ito, Martin, & Nieuwland, 2017; Nieuwland et al., 2018 for DeLong et al., 2005). For the lexical pre-activation effect observed at the pre-nominal gender markers (Otten & Van Berkum, 2009), a replication study observed a pattern (but not a significant effect) consistent with the original study (Kochari & Flecken, 2019). More robust and convincing evidence is therefore needed to understand the predictive language processing that occurs at the pre-nominal linguistic elements.

With regard to the consequence of lexical prediction on processing incoming new words, in addition to the commonly observed N400 effects (associated with semantic processing), a few studies have also reported earlier neural responses such as positivity enhancement around 50–250 ms after the onset of the target word (over posterior/central-posterior electrodes in, e.g., Kim & Lai, 2012; Van Berkum et al., 2005), N1 reduction around 100 ms post-word onset for highly predictable words (Lee, Liu, & Tsai, 2012; Federmeier & Kutas, 2001; Penolazzi, Hauk, & Pulvermuller, 2007), or phonological mismatch negativity reduction around 200 ms post-word onset of expected words (e.g., Hagoort & Brown, 2000). Worth noting is that it is within the N1 window latency (110–130 ms or 120–150 ms post-word onset) that word cloze probability (semantic context) has been found to interact with word frequency or word length, and the N1 response decreases in the condition of high cloze probability with high-frequency or short target words (Lee et al., 2012; Penolazzi et al., 2007). Thus, these early

ERP responses suggest that predictions generated from a semantic context may impact early sensory processes or initial form-related stages of word recognition. However, Nieuwland (2019) provided a dedicated and in-depth review of such "early" effects, and concluded that the reported evidence for such early/low-level facilitating effects is weak or inconsistent, calling for replication of the existing findings.

One goal of the present study was to bring in more data from spoken sentence comprehension to shed further light on lexical predictive processing. Following the literature on lexical prediction, we aimed to investigate anticipatory lexical-semantic processing during natural online speech comprehension by manipulating the semantic constraint of sentence contexts and measuring anticipatory processing of the target nouns prior to their acoustic onset. Furthermore, we were interested in the different processing stages (i.e., pre-semantic acoustic/phonological processing vs. semantic processing) at which lexical predictability facilitates the comprehension of incoming new words in a continuous speech stream.

Unlike existing studies, we investigated lexical prediction by measuring ERP responses at a different pre-nominal linguistic element, namely, transitive verbs. In sentences with SVO order (subject-verb-object), the transitive verb is closely related to the incoming target noun, as it needs a noun argument to form a predicate-argument structure. Transitive verbs (e.g., *BOUGHT* in "*To practice calligraphy, my brother BOUGHT an ink brush*"), when embedded in a relatively strong-constraint semantic context, can trigger anticipation of the lexical-semantic information of the incoming object noun *ink brush*. In a word-by-word reading comprehension study, Li and colleagues (2017) examined anticipatory lexical processing by manipulating the semantic constraint of sentence context and measuring ERPs at the transitive verb preceding the target object noun. Their pretest confirmed the role of verbs in triggering prediction of the incoming target nouns by showing that the cloze probability of the target nouns increased significantly, from 42% to 80%, after the presentation of the transitive verbs. Furthermore, the strong-constraint condition (compared to the weak-constraint condition) elicited not only a reduced N400 effect at the nouns but also an enhanced anterior negativity at the preceding verbs, which is interpreted as a reflection of online anticipatory processing of the target nouns (Li, Zhang, Xia, & Swaab, 2017). This finding is in the visual processing domain and has yet to be replicated in the speech processing domain.

In this study, we will not only consolidate Li et al. (2017) findings on the ERP responses at transitive verbs preceding the target nouns but also reveal, via more analyses, the exact nature of the ERP responses at these verbs. The two possible sources include the easiness of integrating target verbs into the preceding context and the easiness of processing target verbs due to anticipating the upcoming target nouns. We planned to simultaneously model these two processes. Given that the anticipatory and integration processes are tightly related, it is important that we regress out potential verb integration effects and tap into the anticipatory effect by examining the correlation between an ERP component at the transitive verbs and the predictability of their upcoming object nouns. This correlation can then serve as a good indicator that listeners use transitive verbs, combined with preceding context and long-term knowledge, to anticipate the incoming target nouns, which, in turn, affects verb processing. To this end, in addition to the lexical-semantic features of the target nouns, it is crucial that their preceding transitive verbs and other syntactic/thematic properties (such as animacy and syntactic role) are controlled to be exactly the same between the strong- and weak-constraint conditions. This would help to ensure that the correlation between target noun predictability and the amplitude of the ERP component at the verbs cannot be simply explained as due to the integration of the verbs into their preceding context or the processing of their related syntactic/thematic features, thereby lending support to the anticipatory nature of lexical-semantic prediction of the following nouns.

Effects of dialectal experience on lexical prediction and cognitive mechanisms underlying predictive lexical processing

Mixed results have been reported on the role of language experience in shaping anticipatory lexical processing in both behavioral studies (e.g., Chambers & Cooke, 2009; Hopp, 2013; Mitsugi & MacWhinney, 2016) and neural processing research with EEG brain responses (e.g., Martin et al., 2013; Foucart, Martin, Moreno, & Costa, 2014; Ito, Corley, Pickering, Martin, & Nieuwland, 2016; Molinaro, Giannelli, Caffarra, & Martin, 2017). For example, non-native Spanish-English bilinguals have been reported to not anticipate the determiner article (Martin et al., 2013) of an incoming word like native English speakers do. However, French-Spanish (Foucart et al., 2014) and Basque-Spanish (Molinaro et al., 2017) bilinguals have been reported to make use of context information to anticipate the gender-marked determiner of an upcoming noun.

These inconsistent findings on lexical pre-activation in non-native language processing may be attributed to different language experience-related factors. Specifically, critical lexical features or rules deployed in the listeners' native language and maintained in the long-term memory (e.g., determiner articles or grammatical gender markers) as well as general language processing proficiency could give rise to different degrees of native vs. non-native contrasts in lexical prediction. Note that these factors bear a close relationship with two cognitive mechanisms that underpin (at least partly) predictive language processing: One is known to operate quickly and rely on automatic activation of semantic associations with little or no effort; the other is known to be cognitively more demanding and involve effortful mental activities such as complex computation. The two systems have been proposed to constitute the basic foundations of human brain inference processing and are believed to operate simultaneously, supporting each other under most circumstances, with one system typically predominant (Evans & Over, 1996; Kahneman, 2011; Klaczynski & Daniel, 2005; Stanovich & West, 2000).

Predictive processing is considered a specific type of inference (e.g., "predictive inference": McKoon & Ratcliff, 1986; "forward inference": Ferreira & Chantavarin, 2018). Recently, Huettig (2015) linked the two cognitive processing systems to different routes of predictive language processing: Type I prediction relies on automatic activation of associated information; Type II prediction relies on effortful meaning computations or generations, which are more resource-demanding (see also Ferreira & Chantavarin, 2018).

The extent to which predictions in language comprehension are effortfully generated or just reflect an automatic pre-activation of semantic content from a representation of the context is still debated (e.g., Baggio, 2018). Existing studies on non-native predictive processing suggest potential connections to these two mechanisms but have not clearly answered this question, as the long-term knowledge system (such as gender marking, which potentially affects both the automatic activation and effortful computation processes) may not be exactly the same in native and non-native language processing. In contrast to second language experience, dialectal language experience allows listeners with different dialectal backgrounds to react to the same lexical-semantic stimuli in their native language, providing a better controlled experimental setup to tease apart the two possible cognitive mechanisms of predictive lexical processing.

It is no exaggeration that many speakers of the world's languages encounter regional dialectal variations in everyday life. Bidialectism (or multidialectism) may be considered a specific form of bilingualism (or multilingualism) (Albareda-Castellot, Pons, & Sebastián-Gallés, 2011) in which speakers typically have identical or very similar syntactic structures, morphological forms, and semantic representations, but divergent phonology. Dialectal background has been shown to directly influence listeners' perceptual normalization of speech (e.g., Adank, Evans, Stuart-Smith, & Scott, 2009; Floccia, Goslin, Girard, & Konopczynski, 2006; see Samuel & Kraljic, 2009; Alejandrina et al.,

2012 for detailed reviews). It can also affect the mapping between phonological and lexical-semantic levels of representations during word identification (Adank, Van Hout, & Velde, 2007; Sumner & Samuel, 2009 for adult listeners; Floccia, Delle Luche, Durrant, Butler, & Goslin, 2012; Van Heugten & Johnson, 2014 for infants). Some studies have also found that regional dialectal differences introduce difficulty in online speech processing, as indicated by the involvement of domain-general executive control networks (dorsal attentional and default mode networks related to cognitive demands) (Hernández, Ventura-Campos, Costa, Miropadilla, & Avila, 2019) as well as reduced processing speed and increased processing cost (e.g., Floccia et al., 2006; for review see Alejandrina et al., 2012).

When listening to the same stimuli spoken in different dialects, the contextual semantic information that listeners receive and their long-term knowledge system are expected to be the same. Such information is critical for listeners to automatically activate semantic associations or conduct other cognitive-demanding semantic computations for speech comprehension. What is potentially different is the efficiency of bottom-up speech perception (e.g., speech normalization and sound-meaning mapping), which could in turn affect the cognitive resources available to conduct cognitive-demanding meaning computations. Dialectal experience, therefore, provides us an opportunity to examine the two possible cognitive mechanisms of predictive lexical processing (namely, automatic pre-activation of semantic associations and effortful meaning computation) by keeping listeners' long-term mental representation of semantic context constant while changing the processing costs, and consequently, cognitive resources available for meaning computations.

In addition, in the case of home dialect, we know there is a strong link between acoustic/phonological and lexical-semantic levels of representation. If the lexical-semantic content is predicted in advance during language comprehension, this pre-activation is more likely to quickly spread to a low-level form representation due to the facilitated (or more efficient) mapping between semantic and phonological representations. Therefore, echoing the first goal of the present study mentioned earlier, the home dialect provides us with a greater opportunity to observe the early facilitating effect of semantic context on processing new stimuli in continuous speech.

Thus far, two behavioral studies have shown that the semantic context benefit (as indicated by an increased identification rate of highly predictable words compared to lowly predictable ones) is greater for more familiar non-native dialects than for less familiar dialects (Clopper, 2012; Labov & Ash, 1997). What remains open is the earliness of this facilitating effect in online processing and how exactly this context benefit is related to anticipatory lexical processing. The second goal of the present study was, therefore, to examine the cognitive mechanisms (automatic activation of semantic associations or effortful meaning computation) of predictive lexical processing during online language comprehension, by looking at how different dialectal experience (i.e., native, familiar, and unfamiliar) modulates the context-induced predictive processing.

The current study

Mandarin Chinese dialects were used to examine the above issues because the majority of Mandarin Chinese speakers in the People's Republic of China nowadays speak not only their native, early-acquired variety of Mandarin dialects, but at school, they also learn Standard Chinese, which is based mainly on Beijing Mandarin. Furthermore, Mandarin dialects are known for their mutual intelligibility, and speakers can typically communicate with speakers of other Mandarin varieties without serious comprehension problems.

We manipulated the semantic constraint of the sentence context so that the critical noun in a stimulus sentence is either strongly or weakly predictable (i.e., in a strongly or weakly constraining condition). ERPs were measured at both the critical nouns and their preceding transitive verbs. Meanwhile, three different dialect-listening experiences were

included: the early-acquired native dialect (hereafter home dialect); a later-acquired regional dialect with which the listeners are highly familiar (hereafter familiar [FA] regional dialect); and a mutually intelligible but unfamiliar regional dialect which the listeners have only rarely encountered or have not been previously exposed to (hereafter unfamiliar [UN] regional dialect).

It is important to note that we set Beijing Mandarin as the FA regional dialect (in Experiment 1). It is the dialect that Standard Chinese, the national language of China, is based upon (e.g., Hu, 1992; Chirkova & Chen, 2017). While Beijing Mandarin has a few additional features such as more rhotic suffixations and frequent slur in connected speech, the general consensus is that > 90% of the phonological features are shared between the two varieties. Furthermore, many speakers of Beijing Mandarin (including the speaker of the present study) do not speak the dialect with those additional features; their speech is essentially the same as Standard Chinese. Across the country, Standard Chinese is taught in primary school and used in the media and as the medium of instruction in the educational system. For many Chinese, it is the lingual franca for daily activities. Thus, proficient speakers of Standard Chinese are very familiar with Beijing Mandarin, particularly the variety used by our speaker who produced the experimental stimuli.

Among the three dialects, the FA regional dialect (namely Beijing Mandarin) was expected to be fully understood by our participants of the present study. What differs between FA and home dialect is in the age of language acquisition. Home dialect is the one our participants have been extensively exposed to and acquired since their early stage of individual development. Given that early language experience has a maximal effect on ultimate linguistic proficiency (e.g., Lenneberg, 1967; Newman, Bavelier, Corina, Jezzard, & Neville, 2002), the home dialect was expected to have the highest level of perception processing proficiency, relative to the FA and UN regional dialects. The processing proficiency in the FA regional dialect was expected to be higher than that in the UN regional dialect (e.g., Clopper, 2012). Note also that UN differs from the home dialect and FA in all four lexical tones (See more details in the Method section). Thus, the lexical recognition of UN may be hampered to some extent, despite that these dialects are mutually intelligible.

As mentioned earlier, we were interested in the effect of contextual predictability on the neural processing of target object nouns and their preceding transitive verbs, and how dialectal experience modulates this effect. For ERPs at the critical nouns, we were interested in the N400 and early N1 components. The N400 component is considered to reflect the easiness of semantic processing and tends to reduce in amplitude for highly predictable words (e.g., Hagoort & Brown, 2000). The early N1 component usually peaks within the first 150 ms after stimulus onset and can be taken as an index of early acoustic/phonological processing of speech input (Näätänen & Picton, 1987; Obleser, Lahiri, & Eulitz, 2004; Parviainen, Helenius, & Salmelin, 2005; Sanders & Neville, 2003). This component has been reported to associate with the facilitating effect of prior knowledge from matching text on the early perceptual processing of the following spoken word (Sohoglu, Peelle, Carlyon, & Davis, 2012).

We expected that upon articulation of the critical nouns, the strongly predictable ones (compared to the weakly predictable ones) would elicit reduced ERP responses (i.e. N400 and/or N1), indicating a facilitating effect of strong semantic context on processing predictable words. We expected comparable magnitude of N400 reduction for the home and FA regional dialect conditions, as these dialects are mutually intelligible and for the FA chosen in this study, though later acquired, our participants have (near-) native proficiency. In the case of listening to the UN regional dialect, this N400 effect may be reduced in magnitude, if UN indeed can't be understood equally well as the home/FA dialects. The N400 effect may also be delayed in time, if the processing of UN is more cognitively demanding and listeners need more time to achieve lexical access and sentence comprehension.

The N1 effect has been more variably observed (as reviewed earlier). If present, we expect to observe it in the home dialect condition

given that top-down lexical pre-activation as well as efficient information flow between phonetic/phonological and lexical representations is most likely in this condition. As to the non-native language, feedback from lexical down to lower level to support phonetic perception has been found to correlate with language proficiency, with relatively proficient nonnative speakers doing it only in specific conditions (e.g., a word spoken with no conflicting acoustic information) and less-proficient nonnative speakers unable to do it (Samuel & Frost, 2015). The chances of observing an N1 effect of contextual predictability in the FA and UN dialects of the present study are, therefore, likely to be successively smaller.

In short, for ERPs at the target nouns, by examining the time characteristics of the ERP reduction effects (starting from the N400 window latency or from the N1 window latency), we would gain additional insights into the stages of information processing (i.e., early acoustic/phonological processing or later semantic processing) at which the facilitating effect of lexical predictability occurs and how dialectal experience modulates these effects.

For ERPs at the transitive verbs preceding the critical nouns, both negative and positive ERP components might be observed. This is because, when examining anticipatory lexical processing by manipulating pre-nominal linguistic elements (e.g., grammatical gender markers), some studies have reported an enhanced negative ERP response for the expectation-mismatch condition (e.g., Li et al., 2017; Otten & Van Berkum, 2009) while others reported an enhanced positive ERP response (e.g., Wicha, Moreno, & Kutas, 2004; Van Berkum et al., 2005), with different latency and scalp distribution patterns. These ERP responses at the critical verb may be associated with both the easiness of integrating the verbs into the preceding context and the easiness of processing given the incoming critical nouns' predictability. We hypothesized that if listeners do generate hypothesized representations of the incoming nouns, a significant ERP effect should be observed between the strongly and weakly constraining conditions. Moreover, the ERP effect at the verbs should correlate significantly with the predictability of the incoming critical nouns (after excluding the contribution of the easiness of verb integration to this ERP effect), regardless of whether it is a positive or negative effect.

With regard to the way dialectal experience may modulate the target nouns' contextual predictability effects at the pre-nominal transitive verbs, the specific issue we were interested in is whether anticipatory lexical processing is associated with effortful mental activities. If anticipatory lexical processing comes purely from automatic activation of semantic associations, the anticipation effect is expected to be of comparable magnitudes between the home dialect and the FA regional dialect. Given our participants' high-level proficiency in the FA, it is unlikely that this anticipation effect is observed in the home dialect, but completely disappears in the FA regional dialect. Listeners are expected to be able to successfully achieve lexical access and sentence comprehension when listening to FA. In the case of listening to the UN regional dialect, the anticipation effect may be moderated or disappear (compared to the home dialect), due to some potential difficulties in lexical access and sentence comprehension.

If lexical prediction relies on (at least partially) effortful meaning computations in the semantic context, the lexical anticipation effect is likely to be reduced or even disappear in the two regional dialects (compared to the home dialect), given the relatively higher processing cost in non-native dialects and consequently, enhanced cognitive load during perceptual normalization and the mapping between the lower sensory/perceptual and higher semantic/conceptual levels of representations. Note that in the case of UN, there may be additional cost due to the difficulty in lexical access and sentence comprehension.

Two experiments were carried out in the present study. Both used the same set of stimuli mentioned above and differed only in the home dialect of the listeners. As will become clear below, results from Experiments 1 and 2 served to replicate and validate each other's findings.

Experiment 1

Material and methods

Ethics statement

This research was approved by the Ethics Committee of Institute of Psychology, Chinese Academy of Sciences. All participants were over 18 years of age and gave written informed consent. They were notified that their participation was completely voluntary and that they could secede at any time.

Participants

In Experiment 1, twenty-nine right-handed university students (eleven males), all of whom were born and grew up in the city of Dalian (China), participated in this experiment. Their mean age was 22.6 years (range 19–27) at the time of the experiment. None reported any medical, neurological, or psychiatric illness, and all gave informed consent. The data of five participants (two males) were removed from analysis because of excessive EEG artifacts. Therefore, data of twenty-four subjects were included in the final statistical analysis. Participants in Experiment 1 were all native speakers of Dalian Mandarin.

Material

Manipulation of semantic constraint

Mandarin Chinese sentences were used as stimuli. Two manipulations were crucial. First, lexical predictability was manipulated by changing the semantic constraint of the sentence context, as illustrated in Table 1. One hundred twenty-six pairs of sentence frames were constructed; each pair included a strongly constraining sentence frame (e.g., *To practice calligraphy, my brother bought brand-name …*) and a weakly constraining sentence frame (e.g., *To give someone a gift, my brother bought brand-name …*). The last two words of each sentence frame were a transitive verb followed by a classifier/adjective (e.g., *… bought brand-name …*), which were the same between the strongly and weakly constraining conditions, and this transitive verb was the critical verb. Each sentence frame was continued by its best completion, namely the critical noun (e.g., *ink brushes* in the strongly constraining condition and *cigarettes* in the weakly constraining condition) and closed by the same constituent with two/three words. In short, the sentences' semantic context was either strongly or weakly constraining (strongly constraining vs. weakly constraining) and each sentence included a critical transitive verb and a critical noun which serves as the object of the critical verb (see Table 1).

These critical nouns had strong or weak lexical predictability given the preceding context. The critical verbs, when combined with their preceding semantic context, played an important role in triggering the prediction of the lexical meaning of the critical nouns (for detailed reasons see Li et al., 2017). The classifiers/adjectives were carefully selected so that they were statistically equally congruent with the preceding strong- or weak-constraint sentence frames and equally congruent with the following strongly and weakly predictable nouns.

Therefore, the period from the onset of the critical verb to the onset of the critical noun provides the opportunity to examine anticipatory processing of the forthcoming critical noun; the period starting from the onset of the critical noun reflects the integration of the actually presented noun into the preceding context.

Manipulation of dialect experience

The second important manipulation of this study was listeners' dialect experience. This was achieved by recording the written stimulus sentences in three different Mandarin dialects: Dalian Mandarin, Beijing Mandarin, and Jiaozuo Mandarin. Three speakers, all females who were born and grew up in Dalian, Beijing, and Jiaozuo, respectively, produced the same set of stimuli in their own dialects. All three varieties belong to the Mandarin Chinese dialect family. Cognates of the three Mandarin varieties are similar, segmentally speaking, and with regard to lexical tones, they are known to have evolved to a large extent from the same tonal categories historically but differ in the pitch contours of the lexical tones (see Fig. 1). Note that despite the tonal variation at word level, the stimuli are comprehensible at the sentence level.

The recordings were done at a sampling rate of 22050 Hz in a sound-proofed room. To control the potential acoustic differences between the strong- and weak-constraint conditions, the two sentences in each pair (strong vs. weak) were arranged to be read sequentially; the strong-constraint sentence preceded the weak-constraint one in half of the sentence pairs, and the weak-constraint sentence preceded the strong-constraint one in the other half. Speakers were asked to read the sentences aloud at a constant rate without emphasizing/contrasting any particular part of the stimulus sentences.

The acoustic onsets of the critical verbs and critical nouns were defined by two trained native speakers of Standard Mandarin Chinese using the Praat software (version 5.1.27, 1992–2010, Paul Boersma & Davis Weenink). The acoustic onset of each target word was defined as the onset of its first phoneme. Note that when a target word starts with a stop or an affricate (in which case there is an interval of silence due to the oral closure), we marked the onset of the target word at the release of the closure for the phoneme (since there is only silence during the closure and the release serves as a better acoustic landmark for the first phoneme of the target word). For EEG data, the acoustic onsets of the critical nouns were marked during online EEG recording. The acoustic onsets of the critical verbs, however, were calculated and shifted during offline data analysis based on the duration between the onsets of these verbs and nouns.

It is important to note that Dalian Mandarin is the home dialect for participants in Experiment 1; they have been exposed to this dialect since birth, as they were born and grew up in the city of Dalian; their parents and everyone in the local community speak Dalian Mandarin on a daily base. The participants are familiar with Beijing Mandarin because, as mentioned in the introduction, it is essentially the same as Standard Chinese, which is taught in primary school and used in the media. Jiaozuo Mandarin, however, although intelligible, is unfamiliar to our participants, as it is rarely encountered by these participants and has significant tonal differences from Dalian Mandarin. In short, in terms of the age of language acquisition, the participants have only

Table 1
Illustration of the experimental materials in the strongly vs. weakly constraining conditions.

Conditions	Example sentences
Strong	哥哥 为了 练习 书法 "买了" 老牌的 毛笔 带回家。
	Brother/to /practice/calligraphy/"buy"/brand-name/ink brush/take them home
	To practice calligraphy, my brother "bought" *brand-name* ink brushes and took them home.
Weak	哥哥 为了 送人 礼物 "买了" 老牌的 香烟 带回家。
	Brother/to /give someone/gift /"buy"/brand-name/cigarettes/take them home
	To give someone a gift, my brother "bought" *brand-name* cigarettes and took them home.

Note. The underlined words are the critical nouns; the italic words are the adjectives/classifiers immediately preceding the critical nouns; the words in quotation marks are the critical verbs immediately preceding the adjectives/classifiers.
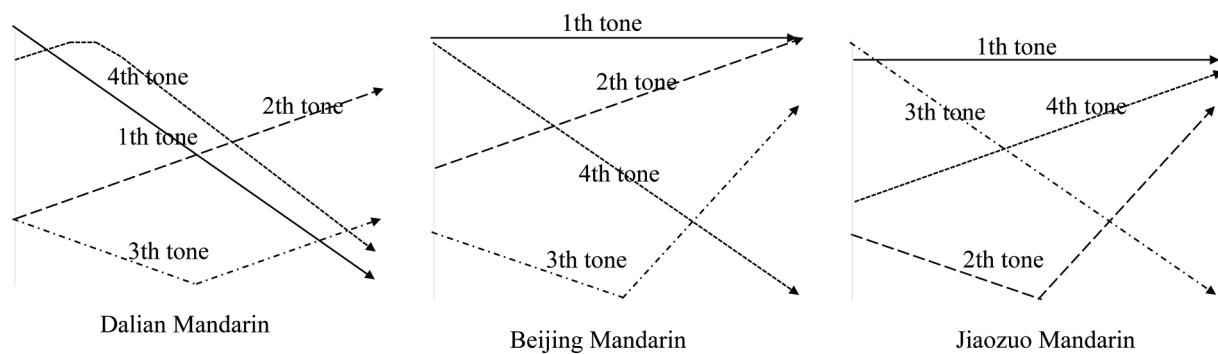
Fig. 1. Schematic representations of the lexical tone pitch contours of the three Mandarin dialects. In Beijing Mandarin, T1 has a high-level pitch contour, T2 a rising contour, T3 a low dipping contour, and T4 a falling contour. In Dalian Mandarin, Tone1 differs from that of Beijing Mandarin: It has a falling contour like T4 but with a more abrupt/steeper falling contour than T4. The two dialects have similar pitch contours for the other three lexical tones. The tonal pitch contours of Jiaozuo Mandarin, however, differ more drastically from that of Dalian Mandarin. Its T1 has a high-level pitch contour (like Beijing Mandarin but differs from Dalian Mandarin); T2 has a dipping pitch contour, T3 a falling contour and T4 a rising contour, which are different from both Dalian Mandarin and Beijing Mandarin. In short, there are more similarities between the tonal pitch contours of Beijing Mandarin and Dalian Mandarin than that between Dalian Mandarin and Jiaozuo Mandarin.

been extensively exposed to and used Dalian Mandarin since their early stage of individual development. In terms of familiarity, both Dalian Mandarin and Beijing Mandarin are familiar to the listeners, while Jiaozuo Mandarin is unfamiliar. As mentioned in the introduction section, we expected a higher level of processing proficiency in the home dialect (namely, Dalian Mandarin in Experiment I) compared to that in the FA and UN regional dialects; the processing proficiency in the FA regional dialect was in turn expected to be higher than the UN regional dialect. It is also important to bear in mind that both the home dialect and the FA regional dialect were expected to be fully understood by the participants.

Taken together, the design is full factorial with all combinations of the factors Semantic Constraint (strong vs. weak) and Dialect (home, FA regional, vs. UN regional). Moreover, we measured the semantic constraint effect at both the critical object nouns and the preceding critical transitive verbs.

Written pre-tests for stimuli verification

To validate the degree of critical noun predictability given the sentence contexts, we conducted a series of cloze probability tests by visually presenting the sentence frames until (1) the word before the critical verb position (cloze1), (2) the word before the classifier/adjective position (cloze2), or (3) the word before the critical noun position (cloze3). Twenty-four subjects participated in the cloze1 test, twenty-four new participated in the cloze2 test, and another set of thirty-six new subjects participated in the cloze3 test. None of these participants attended the EEG experiment. They were instructed to fill in the first event that came to their mind and complete each sentence in a meaningful way. In the strongly constraining condition, the cloze probabilities of the critical nouns were 42%, 81%, and 77% for cloze1, cloze2, and cloze3, respectively; in the weakly constraining conditions, the cloze probabilities of the critical nouns were 18%, 24%, and 27%, respectively (see Table 2). The ANOVA (analyses of variance) with Semantic Constraint (strong vs. weak) and Position (cloze1, cloze2, vs. cloze3) as independent factors showed a significant main effect of Semantic Constraint ($F_{(1,125)} = 930.7$, $p < .0001$) and a significant interaction between Semantic Constraint and Position ($F_{(2,250)} = 96.86$, $p < .0001$). Further simple analysis showed that the critical nouns' cloze probability in the strongly constraining condition was significantly higher than that in the weakly constraining condition for all three pre-test positions (cloze1: $F_{(1,125)} = 83.38$, $p < .0001$; cloze2: $F_{(1,125)} = 1541.31$, $p < .0001$; cloze3: $F_{(1,125)} = 811.43$, $p < .0001$). More importantly, ANOVAs with the difference-score (strong-constraint minus weak-constraint) as a dependent factor also revealed a significant

effect of Position ($F_{(2,250)} = 96.86$, p $< .0001$), suggesting that target noun predictability increased greatly after the critical verb was presented (with Bonferroni correction: p $< .0001$ for both cloze1 vs. cloze2 and cloze1 vs. cloze3 comparisons). These pre-tests ensured that our manipulation of strongly and weakly constraining contexts was successful, and the critical verbs (combined with preceding semantic contexts) played an important role in triggering the prediction of the forthcoming critical nouns' lexical meaning.

To further examine the semantic congruency between the critical nouns and their sentence context, twenty-four participants who did not attend the EEG experiment and other pre-tests were visually presented with the frame sentences until the critical nouns and instructed to rate, on a 7-point scale (from -3 to 3), the degree of the nouns' semantic congruency with their corresponding sentence contexts. Negative numbers indicate levels of incongruence and positive numbers levels of congruence, with zero being neutral. The larger the scores, the more congruent these words are with the preceding context. Results of the critical nouns' semantic congruency are presented in Table 2.

We also examined the semantic congruency of the critical verbs and classifiers/adjectives (immediately following the critical verbs) with their preceding sentence context, respectively. Similar to the pre-test for critical nouns, twenty-four new participants were visually presented with the stimulus sentences until the critical verbs, and another twenty-four participants were visually presented with stimulus sentences until the classifiers/adjectives. None of these participants attended the EEG experiment and other pre-tests. They were instructed to rate the degree of these verbs' (or classifiers/adjectives') semantic congruency given their preceding context on a 7-point scale (from -3 to 3). In addition, we also calculated the cloze probability of the critical verbs based on the participants' responses in the cloze1 pre-test mentioned above. Results of the verb cloze rating and congruency of verbs are presented in Table 2.

As the critical nouns were all the best completion in their semantic context, they are not the same words in the strongly and weakly constraining conditions. We further examined the degree of semantic congruency between the critical nouns and the preceding classifiers/adjectives, as well as these nouns' word frequency and number of strokes. Like the congruency pre-test mentioned above, twenty-four participants who did not attend the EEG experiment or other pre-tests were presented with adjective/classifier-noun pairs and instructed to rate the degree of congruency for each pair on a 7-point scale (from -3 to 3). Detailed results on the noun-adjective/classifier congruency, frequencies of nouns, and the number of strokes of the nouns are also reported in Table 2.

Table 2
Properties of the critical words in the strong- and weak-constraint conditions.

	Strong-constraint		Weak-constraint	
	Mean	SD	Mean	SD
Cloze1	42.1%	27.9%	18.1%	12.0%
Cloze2	80.6%	15.2%	24.1%	10.7%
Cloze3	77.3%	14.6%	27.2%	13.6%
Verb cloze	31.7%	33.4%	16.5%	27.7%
Congruency of verbs	2.40	0.62	2.38	0.58
Congruency of adjectives/classifiers	2.53	0.48	2.50	0.54
Congruency of nouns	2.77	0.28	2.43	0.58
Noun–adjective/classifier congruency	2.40	0.75	2.42	0.59
Frequency of nouns	12.52	9.40	14.16	18.12
Number of strokes of nouns	16.6	4.9	15.9	5.9

Note: Cloze1, Cloze2, and Cloze3 indicate the cloze probabilities of the critical nouns when the sentence frame was presented until the word before the critical verb, until the critical verb, and until the classifier/adjective, respectively. Verb cloze indicates the cloze probability of the critical verbs when the sentences were presented until the word before these verbs. 'Noun–adjective/classifier congruency' indicates the degree of congruency between the critical nouns and their preceding adjectives/classifiers.

Modification and verification of acoustic stimuli

Note that all pre-tests were conducted by using written versions of the stimulus sentences. To control the potential confounding factors coming from the speech stimuli (in particular from the acoustic parameters of the critical nouns), a few further steps were taken. First, the average sentence intensity of each sentence was scaled to 56 dB. The average durations of the stimulus sentences are comparable, as the ANOVAs with Semantic constraint and Dialect as independent factors showed no significant main effect nor interaction ($ps > 0.7$): Beijing-Strong (3426 ms; SD = 351), Beijing-Weak (3421 ms; SD = 443), Dalian-Strong (3425 ms; SD = 360), Dalian-Weak (3432 ms; SD = 361), Jiaozuo-Strong (3431 ms; SD = 364), and Jiaozuo-Weak (3437 ms; SD = 376).

We further analyzed the acoustic parameters of the critical words in the six experimental conditions: (1) duration, F0 pitch maximum, and intensity of the two words preceding the critical nouns (i.e. the critical verbs and adjectives/classifiers); (2) duration, F0 pitch maximum, and intensity of the critical verbs; and (3) duration, F0 pitch maximum, and intensity of the critical nouns. For details of the above acoustic parameters, please see Table 3. Impressionistically speaking, the results suggested no particular pattern that would bias the processing of the stimuli for our participants. Nevertheless, the results were modelled further in the linear mixed-effects analyses of the EEG data to examine whether, after regressing out the effects of potentially confounding factors, the effects of factors of our interest (i.e. the semantic constraint and dialect experience) would remain significant.

In total, we included 126 groups of experimental stimulus sentences; each group included six conditions (with all combinations of the

Semantic Constraint and Dialect factors). The experimental materials were grouped into three lists of 252 sentences according to the Latin square procedure based on the three dialect conditions (home, FA regional, vs. UN regional); the strong-constraint and weak-constraint sentences from the same stimulus group were included in the same list. In each list, there were an equal number of sentences (forty-two sentences) for each of the six experimental conditions and an additional eighty filler sentences (with twenty-six sentences for each dialect and two additional sentences that were placed at the beginning/end of the material list). Meanwhile, the whole list of sentences (332 sentences including 252 experimental and 80 filler sentences) was divided into four blocks. The presentation order of each pair of sentences (namely, the strong- and weak-constraint sentences coming from the same stimuli pair) was counterbalanced between subjects, and the presentation of the two sentences for each pair was separated by at least one block.

Procedure

A short practice session with eighteen sentences (six sentences for each of three kinds of dialect) was carried out before the EEG experiment. The subjects listened to a sentence (e.g., *It was recently found that watching TV might be one of the reasons that drive the occurrence of dementia*) and then were asked to judge how correctly a visually presented sentence (called a "question sentence" hereafter; e.g., *Dementia might be related to watching TV*) reflected its meaning. If all questions were answered correctly, the real EEG experiment started. If not, they were asked to listen to the sentences again. After the first or the second presentation of the sentences, all participants answered with 100% accuracy.

During the main EEG experiment, subjects were asked to listen to each sentence for comprehension while their EEG signals were recorded. After a short practice session that consisted of ten sentences, the trials were presented in the above-mentioned four blocks, each lasting approximately eleven minutes and separated by a brief resting period. Each trial consisted of a 300 ms auditory warning tone, followed by 700 ms of silence and then the target sentence. To inform subjects when to fixate and sit still for the EEG recording, an asterisk was displayed from 500 ms before the onset of the sentence until 1000 ms after the offset. As in the short practice period, the participants listened to each sentence for comprehension and were then asked to judge the correctness of the question sentence (18 sentences for each dialect or 54 per block) regarding its meaning. Different from the practice session, participants only listened to the sentences once.

EEG acquisition

The EEG was recorded (0.05–100 Hz, sampling rate 500 Hz) from 64 Ag/AgCl electrodes mounted on an elastic cap, with an online reference linked to the left mastoid and an offline algebraic re-reference linked to the left and right mastoids. EEG and EOG data were amplified

Table 3
The acoustic parameters (including duration 'ms', f0 pitch maxmum 'Hz', and intensity 'dB') of the critical words of the experimental sentences.

		Beijing		Dalian		Jiaozuo	
		StrongM(SD)	WeakM(SD)	StrongM(SD)	WeakM(SD)	StrongM(SD)	WeakM(SD)
Verb + adj	duration	708 (145)	717 (154)	792 (156)	799 (156)	755 (144)	761 (141)
	pitch	295 (40)	294 (48)	303 (32)	311 (49)	238 (59)	237 (52)
	intensity	56.1 (1.1)	56.1 (1.2)	55.6 (1.2)	55.7 (1.3)	56.2 (1.1)	56.1 (1.1)
Verb	duration	347 (84)	368 (181)	383 (88)	382 (92)	387 (93)	388 (95)
	pitch	278 (46)	275 (55)	281 (29)	280 (37)	229 (15)	227 (18)
	intensity	56.0 (1.7)	56.1 (1.8)	56.2 (1.5)	56.4 (1.7)	56.6 (1.4)	56.6 (1.5)
Noun	duration	466 (50)	468 (60)	424 (47)	417 (50)	425 (54)	417 (57)
	pitch	271 (55)	267 (54)	290 (40)	280 (32)	234 (29)	233 (26)
	intensity	55.6 (1.5)	55.7 (1.7)	55.8(1.8)	55.6 (1.9)	55.3 (1.9)	55.2 (1.7)

Note: 'verb + adj' indicates the two words (the critical verb and adjective/classifier) preceding the critical noun.

with AC amplifiers (Synamps, Neuroscan Inc.). Vertical and horizontal electro-oculargrams were monitored. All electrode impedance levels (EEG and EOG) were kept below 5 kΩ.

ERP analysis

To investigate the integration processing of critical nouns, analyses were time-locked to the onset of the nouns. To examine the potential anticipatory processing of critical nouns prior to their acoustic onset, analyses were time-locked to the onset of the critical verbs. We analyzed how the ERP amplitude at the verbs was associated with the easiness of integrating these verbs into the preceding context and the forthcoming critical nouns' predictability.

The Fieldtrip toolbox for MATLAB (Oostenveld, Fries, Maris, & Schoffelen, 2011) was used to preprocess the EEG data and calculate ERPs. The raw EEG data were first filtered with a band-pass filter 0.1–40 Hz. Then, ocular artifacts were isolated using independent component analysis and removed by projecting the data onto the pseudoinverse of the artifact components. These ocular-artifact-corrected data were divided into epochs ranging from 100 ms before the onset of the critical words to 1000 ms after the onset of these words. A time window of 100 ms preceding the onset of the critical words was used for baseline correction. Subsequently, trials contaminated by muscle artifacts, electrode drifting, amplifier saturation, or other artifacts were identified by semiautomatic artifact rejection (automatic criterion: signal amplitude exceeding ± 75 μV, followed by a manual check). Trials containing artifacts mentioned above were rejected. The ANOVA test with conditions (six levels) as independent factors showed that there was no significant difference in the rejected trials from the six experimental conditions at both nouns ($F_{(5,115)} = 1.38$, $p = .249$) (8.7% [SD = 5.4], 6.6% [SD = 4.6], 9.0% [SD = 7.0], 7.0% [SD = 6.6], 8.8% [SD = 6.1], and 8.0% [SD = 7.7] for Beijing-Strong, Beijing-Weak, Dalian-Strong, Dalian-Weak, Jiaozuo-Strong, and Jiaozuo-Weak conditions, respectively) and verbs ($F_{(5,115)} = 1.20$, $p = .315$) (8.0% [SD = 6.8], 8.5% [SD = 5.3], 10.6% [SD = 7.3], 9.6% [SD = 7.0], 9.8% [SD = 6.4], and 9.6% [SD = 8.4] for Beijing-Strong, Beijing-Weak, Dalian-Strong, Dalian-Weak, Jiaozuo-Strong, and Jiaozuo-Weak conditions, respectively). Finally, averages were computed for each participant, each condition, and at each electrode site before grand averages were calculated across all participants.

Statistical analysis

To examine whether the manipulation of Semantic Constraint (strong vs. weak) elicited a significant ERP effect and how it interacted with Dialect (home, FA regional, vs. UN regional), a cluster-based random permutation test was conducted within 0–600 ms post-word onset (in steps of 2 ms) over forty-seven electrodes. This permutation test (implemented in FieldTrip, Maris & Oostenveld, 2007) computes a separate *t*-test at every time point for every electrode site and uses a cluster method to correct for multiple comparisons. It is a good option to control the overall Type I error rate when no a priori information is available to guide the choice of latency windows and electrode sites (Luck & Gaspelin, 2017).

Focusing on the window latencies and scalp regions that showed a significant main (or simple main) effect of Semantic Constraint in the permutation test, we then performed linear mixed-effects model analyses (Baayen, Davidson, & Bates, 2008) of the signal trial ERP data, using the "lmerTest" package (Kuznetsova, Brockhoff, & Christensen, 2017) in the R software (R CoreTeam, 2018). In contrast to the cluster-based test, this approach simultaneously models the variance associated with each subject and each item. In particular, when analyzing the effects of continuous predictor variables such as the cloze probability of target nouns, linear mixed-effects regression is able to tease apart each factor's contributions. The linear mixed-effects model analyses were conducted in two steps: first categorical analysis and then continuous

analysis. Both types of analysis were conducted separately for verbs and nouns.

The categorical analyses were conducted first to facilitate model convergence and to follow the present experiment design. Two factors Semantic Constraint (strong vs. weak) and Dialect (home, FA regional, vs. UN regional) were modeled as categorical predictors on both the target nouns and their preceding transitive verbs. This type of analysis aimed to further examine the effect of Semantic Constraint and its interaction with Dialect by simultaneously modeling other potentially confounding factors (e.g., word frequency, number of strokes of nouns, and acoustic parameters) and the variance associated with each subject and each item.

If the main or simple main effect of Semantic Constraint reached significance in the above analyses, a further continuous analysis of single-trial data was conducted to examine the exact reasons for this Semantic Constraint effect at the target nouns and the preceding transitive verbs using continuous independent variables (obtained in the pretests). At the verbs, the factor Semantic Constraint was closely associated with (1) the verbs' semantic congruency (indicated by the values of the verb congruency pretest), (2) the preceding context (indicated by the values of the cloze1 pretest and the values of verb cloze probability), and (3) the predictability of forthcoming critical nouns after the presence of these verbs (indicated by the values of the cloze2 pretest). Here, the first factor is related to the easiness of integrating the verbs into the preceding context while the last related to the anticipatory processing of incoming critical nouns. At the target nouns, the factor Semantic Constraint was closely related to both the nouns' predictability (indicated by the values of cloze3) and semantic congruency (indicated by the values of the noun congruency pretest).

Note that the values of cloze2 indicate the critical nouns' predictability right after the critical verbs have been presented but before the following adjectives/classifiers are presented. Therefore, this measurement is more informative for the analysis at the verbs. The values of cloze3 indicate the cloze probability of the critical nouns after the adjectives/classifiers and before the nouns have been presented. This measurement is more informative for the analysis at the nouns.

In summary, to examine the exact nature of the semantic constraint effect, further continuous analyses of the single-trial ERP data (linear mixed-effects model analysis) were conducted. The above factors that were closely related to semantic constraint and other potentially confounding factors were simultaneously modeled as continuous predictors. If listeners indeed conduct anticipatory processing of the following nouns after hearing the verbs, their ERP responses at the verbs should then correlate well with the values of the cloze2 pretest after the contributions of other factors have been considered. Detailed statistical analyses are described below.

Further details of the cluster-based random permutation tests

The cluster-based random permutation tests (0–600 ms from post-word onset; in steps of 2 ms) were performed over forty-seven electrodes (F7/F8, F5/F6, F3/F4, F1/F2, FZ, FC5/FC6, FC3/FC4, FC1/FC2, FCZ, C5/C6, C3C4, C2/C1, CZ, CP5/CP6, CP3CP4, CP2/CP1, CPZ, P7/P8, P5/P6, P3/P4, P1/P2, PZ, PO7/PO8, PO5/PO6, PO3/PO4, POZ, OZ). The other electrodes were removed from the statistical analysis in the present study because they were more likely to be affected by artifacts. The dependent measure was the ERP amplitude at the specific time point and electrode. For every data point ("electrode by time") of two conditions, a simple dependent-samples *t*-test was performed. All adjacent data points exceeding a preset significance level (0.05%) were grouped into clusters. Cluster-level statistics were calculated by taking the sum of the *t*-values within every cluster. The significance probability of the clusters was calculated by the so-called Monte Carlo method with 1000 random draws.

We first examined whether there was a significant two-way Semantic Constraint × Dialect interaction. That is, we calculated ERPs

for the six experimental conditions separately (two Semantic Constraint by three Dialect). Permutation tests were performed on the differences between the difference waveforms (Home$_{strong}$-minus-Home$_{weak}$ vs. FAregional$_{strong}$-minus-FAregional$_{weak}$; Home$_{strong}$-minus-Home$_{weak}$ vs. UNregional$_{strong}$-minus-UNregional$_{weak}$; FAregional$_{strong}$-minus-FAregional$_{weak}$ vs. UNregional$_{strong}$-minus-UNregional$_{weak}$). If any of these comparisons showed a significant effect (with Bonferroni corrections), the Semantic Constraint × Dialect interaction was considered to be significant. Then, the permutation test was conducted on the difference between the strong- and weak-constraint conditions for each dialect separately. Otherwise, the main effect of Semantic Constraint was analyzed by combining the three dialects.

Further details of the linear mixed-effects models

Using a spatiotemporal region-of-interest approach, linear mixed-effects models were performed with the scalp regions and window latencies selected based on the permutation test results. The dependent measure was the average voltage across the related channels and window latencies.

For the categorical analysis of the single-trial data, we included Semantic Constraint (strong vs. weak) and Dialect (home, FA regional, vs. UN regional) as categorical predictors; their interaction was also entered as a fixed effect. In addition, the z-scored values of the acoustic parameters (F0 pitch maximum, duration, and intensity) were entered into the model as fixed factors. We started with the maximal random effect structures, which were trimmed if the model did not converge. Finally, the following models were used.

For verbs: *eegdata ~ constraint*dialect + zVerbPitch + zVerbDuration + zVerbIntensity + (constraint | subject) + (constraint | item)*

For nouns: *eegdata ~ constraint*dialect + zfrequency + zstroke + zNounPitch + zNounDuration + zNounIntensity + (constraint | subject) + (constraint | item)*

For the continuous analysis of the single-trial data, we first performed a correlation analysis of the factors that were closely related to semantic constraint (including cloze1, verb congruency, verb predictability, and cloze2 for analysis at the critical verbs; including cloze3 and noun congruency for analysis at the critical nouns) using the "cor.test" function package in the R software. For the verbs, this correlation analysis revealed a significant correlation between cloze1 and verb congruency ($r = 0.278$) and between cloze1 and cloze2 ($r = 0.519$); the correlation between verb predictability and cloze1 (cloze2, or verb congruency; $r = 0.381$, $r = 0.275$, $r = 0.346$, respectively) also reached significance. For the nouns, the correlation between cloze3 and noun congruency reached significance ($r = 0.339$). These correlations were not a principled obstacle to our linear mixed-effects model analysis. Variance inflation factors for these continuous predictors were all below 1.42, which is below the value considered problematic due to high multicollinearity (e.g., Zuur, Ieno, & Elphick, 2010).

After the correlation tests, we conducted a linear mixed-effects regression analysis for verbs and nouns separately, in which the z-scored values of the above factors and other potentially confounding factors were entered into the models as continuous predictors. We removed random correlations to facilitate model convergence.

For verbs: *eegdata ~ zcloze2 + zVerbCongruency + zVerbCloze + zcloze1 + zVerbPitch + zVerbDuration + zVerbIntensity + (zcloze2 + zVerbCongruency + zVerbCloze + zcloze1 || subject) + (zcloze2 + zVerbCongruency + zVerbCloze + zcloze1 || item)*

For nouns: *eegdata ~ zcloze3 + zNounCongruency + zfrequency + zstroke + zNounPitch + zNounDuration + zNounIntensity + (zcloze3 + zNounCongruency || subject) + (zcloze3 + zNounCongruency || item)*

We extracted the *F*-value and *p*-value associated with each fixed effect in all of the above models. In addition, for the linear mixed-effects regression models, we extracted a coefficient estimate with a 95% confidence interval (CI) associated with each significant fixed effect. CIs were computed with the Wald method and *p*-value with the normal approximations. All output β estimates and 95% CIs were transformed from z-scores back to raw scores of the corresponding fixed factor (e.g., 0–100% cloze-range for the factors associated with cloze probability).

As the linear mixed-effects model analyses were conducted over multiple region-of-interest scalp locations and time windows, multiple comparisons were corrected to control the false discovery rate, by using the Benjamini and Hochberg (1995) (implemented in R's p.adjust) method. We applied this correction to the *p*-values associated with samples from all region-of-interest channel locations. This correction was applied separately to the early (e.g., N1) and later (e.g., N400) windows because the false discovery rate procedure can be too lenient outside the window latency containing relatively larger ERP effects (Groppe, Urbach, & Kutas, 2011).

Results

Behavioral results

For each participant included in the final statistical analysis of the EEG data, the rate of accuracy in his/her answers to the comprehension questions in each dialect (across all conditions) was always higher than 75%. The averaged accuracy rates for home, FA regional, and UN regional dialects were 94.44% (SD = 6.27), 92.51% (SD = 6.62), and 87.20% (SD = 7.00), respectively. The ERP results are reported below and summarized in Table 4.

ERP results at the critical nouns

Cluster-based random permutation test results

The cluster-based random permutation test revealed a significant interaction between Semantic Constraint and Dialect, as indicated by the significant differences in the difference waveforms (strong-minus-weak) between the home and FA regional dialects ($p = .015$, with a

Table 4
Summary of the ERP effects elicited at the critical transitive verbs and target object nouns in Experiment 1.

	Time window	Description of significant effects	Polarity	Topography
Verbs	120–180 ms	No effect	N.A.	N.A.
	300–500 ms	*Significant interaction (S*D):* Enhanced positivity for strong-constraint in home dialect only (associated with cloze2)	Positivity	Broad distribution with maximum at frontal electrodes
Nouns	120–180 ms	*Significant interaction (S*D):* Reduced N1 for strong-constraint in home dialect only (associated with cloze3)	Negativity	Frontal-central distribution
	300–500 ms	*Significant main effect (S):* Reduced N400 for strong-constraint in all three dialects (associated with cloze3)	Negativity	Broad distribution in home dialect; central-parietal distribution in the two regional dialects

Note. S and D indicate the 'Semantic Constraint' and ''Dialect', respectively.

cluster extending approximately from 160 ms to 210 ms post-noun onset) and between the home and UN regional dialects ($p = .045$, for a time window of approximately 90 ms to 220 ms post-noun onset). The subsequent simple effect analysis showed that in listening to the home dialect (Dalian Mandarin), there was a significant difference between the strong- and weak-constraint conditions ($p < .001$; at an interval of approximately 90 ms to 600 ms post-noun onset); this effect was pronounced over a widespread set of electrodes, including the anterior, central, and posterior electrodes. Under the other two regional dialect conditions, we also found a significant difference between the strong- and weak-constraint conditions (FA regional dialect: $p < .01$, for a time window of approximately 240 ms to 410 ms post-noun onset; UN regional dialect: $p = .042$, for a time window of approximately 260 ms to 350 ms post-noun onset); this semantic constraint effect was most pronounced over the central-parietal electrodes for both regional dialect conditions (see Figs. 2 and 4).

As shown in Fig. 2, in the case of the home dialect, the ERP deflections underlying the sustained semantic constraint effect had two clearly negative peaks; the first in the N1 window latency and the second in the N400 widow latency. Therefore, we classified this semantic constraint ERP effect as a combination of the N1 and N400 effects. In contrast, the semantic constraint effects observed in the cases of the two regional dialects occurred within the window latency of N400 and were underlain by negative ERP deflections, hence classified as the N400 effect.

Linear mixed-effects model analysis results

As the cluster-based random permutation test showed that the Semantic Constraint effect observed in the two regional dialect conditions was most pronounced over the central-parietal electrodes, linear mixed-effects model analyses were performed separately for the frontal (single-trial magnitude averaged across *F5/F6, F3/F4, F1/F2, FZ, FC5/FC6, FC3/FC4, FC1/FC2, FCZ*), central (averaged across *C5/C6, C3C4, C2/C1, CZ, CP5/CP6, CP3CP4, CP2/CP1, CPZ*), and parietal (averaged across *P7/P8, P5/P6, P3/P4, P1/P2, PZ, PO7/PO8, PO5/PO6, PO3/PO4, POZ*) electrodes. These analyses were also conducted for the N1 (120–180 ms) and traditional N400 (300–500 ms) window latencies separately. The window latency of 120–180 ms post-noun onset was selected because the N1 component peaked around 150 ms post-noun onset (by visual inspection). Note that the cluster-based random permutation test, which optimally handles the multiple-comparisons problem, demonstrated that the semantic constraint effect started already approximately from 90 to 140 ms post-noun onset.

Within the N1 window latency, firstly, the categorical analysis of single-trial data revealed a significant interaction between Semantic Constraint and Dialect in the frontal and central scalp regions ($F_{frontal} = 6.95$, $p < .005$; $F_{central} = 3.63$, $p = .040$); in contrast, in the parietal region, neither the main effect of Semantic Constraint nor its interaction with Dialect reached significance (all $ps > .092$). Further simple effect analysis demonstrated that, compared to the weak-constraint condition, the strong-constraint condition elicited a significantly reduced negative deflection in the case of the home dialect ($F_{frontal} = 11.17$, $p < .005$; $F_{central} = 9.45$, $p < .005$) but not in the two regional dialect conditions (all $ps > .540$). Secondly, further continuous analysis of single-trial data (linear mixed-effects regression) found that, in listening to the home dialect, smaller negative deflections (N1) were significantly associated with a larger value of cloze3 (frontal: $\beta = 1.73$, CI [0.76, 2.70], $F = 12.26$, $p < .005$; central: $\beta = 1.68$, CI [0.49, 2.87], $F = 7.68$, $p < .01$) and a larger pitch magnitude of nouns (frontal: $\beta = 0.95$, CI [0.01, 1.88], $F = 3.95$, $p = .048$; central: $\beta = 1.17$, CI [0.30, 2.05], $F = 6.90$, $p = .018$); however, neither the noun congruency (all $ps > .090$) nor the other factors' effects (all $ps > .265$) reached significance.
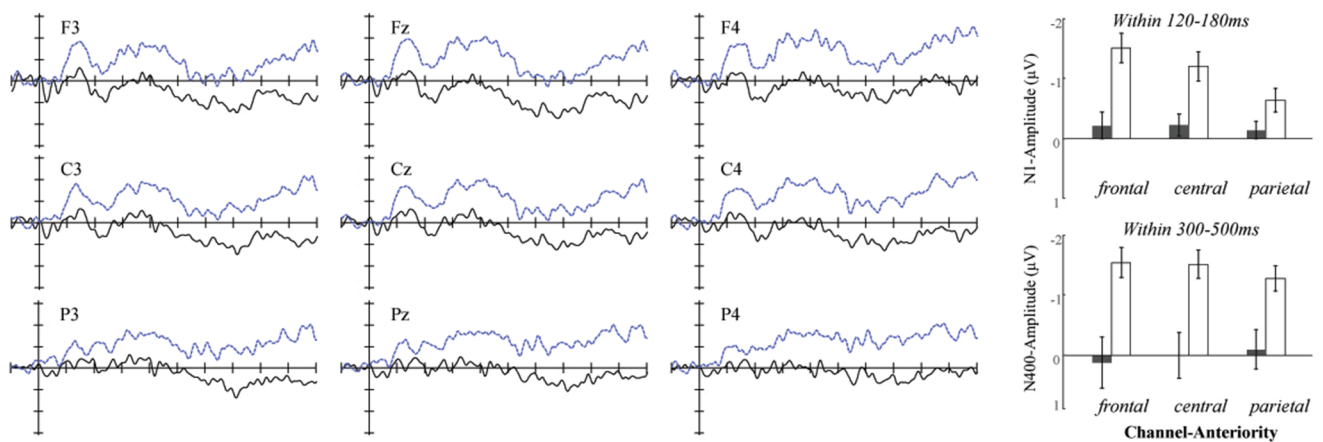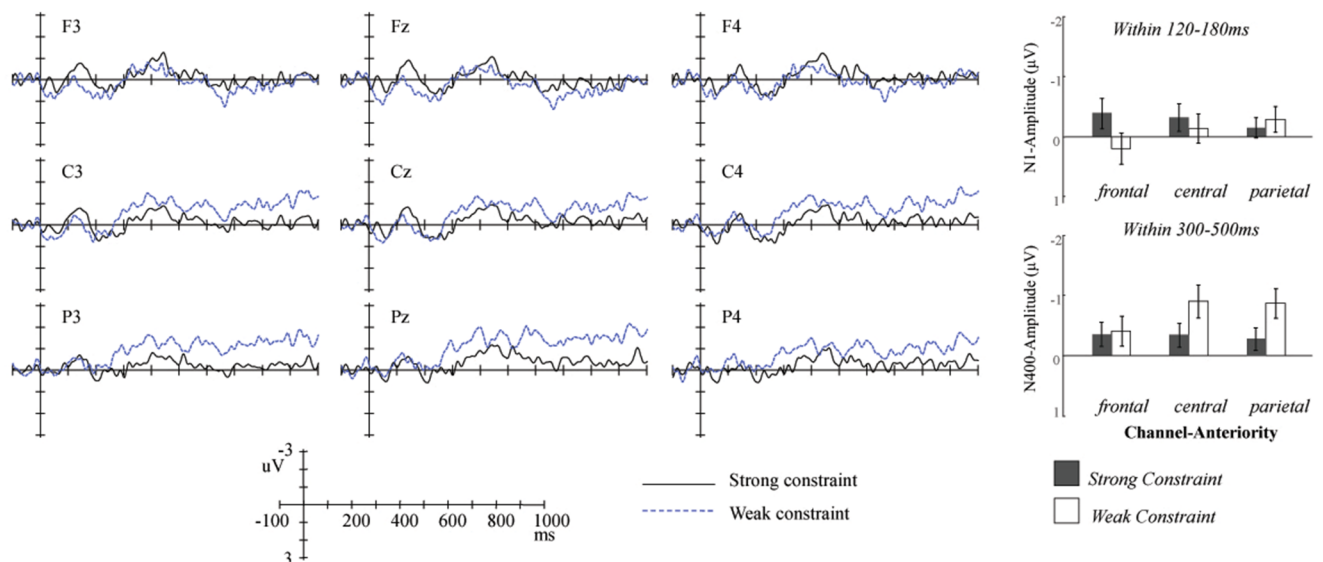
Within the N400 window latency, firstly, the categorical analysis of single-trial data resulted in a significant main effect of Semantic Constraint ($F_{frontal} = 10.49$, $p < .005$; $F_{central} = 17.48$, $p < .001$;

$F_{parietal} = 17.93$, $p < .001$), indicating that the strong-constraint condition elicited a reduced N400 (compared to the weak-constraint condition). The Semantic Constraint × Dialect interaction reached significance only over the frontal electrodes ($F_{frontal} = 4.86$, $p = .023$; $F_{central} = 1.97$, $p = .210$; $F_{parietal} = 0.38$, $p = .685$). A subsequent simple effect analysis over the frontal region showed that at the frontal channels, the Semantic Constraint effect reached significance only in the home dialect condition ($F_{frontal} = 19.39$, $p < .0001$). That is, the context-constraint-related N400 effect had a broad distribution in the home dialect condition and a central-parietal distribution in the two regional dialect conditions. Further continuous analysis of single-trial data showed that, smaller N400 deflections were significantly associated with a larger value of cloze3 (frontal channels for home dialect: $\beta = 2.97$, CI [1.43, 4.50], $F = 14.51$, $p < .001$; central channels for all dialects: $\beta = 1.89$, CI [1.10, 2.66], $F = 22.32$, $p < .001$; parietal channels for all dialects: $\beta = 1.72$, CI [0.98, 2.46], $F = 20.54$, $p < .001$) and a shorter duration of nouns (central: $\beta = -0.44$, CI [-0.82, -0.06], $F = 5.06$, $p = .037$; parietal: $\beta = -0.48$, CI [-0.84, -0.13], $F = 7.12$, $p = .023$); however, neither the noun congruency (all $ps > .430$) nor the other factors' effects (all $ps > .196$) reached significance.

Relative magnitude and onset latency of the N400 effect

To directly compare the magnitudes (effect sizes) of the N400 effect in the three dialect conditions, further categorical analyses of single-trial data were performed over the data of each two dialect conditions separately. These separate analyses showed no significant differences in the context-constraint-related N400 effect between the home and FA conditions, as indicated by the lack of significant interaction between Semantic Constraint and Dialect ($F_{frontal} = 5.14$, $p = .070$; $F_{central} = 3.03$, $p = .122$; $F_{parietal} = 0.42$, $p = .516$); the difference between the FA and UN regional dialect conditions also did not reach significance, ($F_{frontal} = 0.44$, $p = .898$; $F_{central} = 0.02$, $p = .898$; $F_{parietal} = 0.02$, $p = .898$). The N400-effect-magnitude differences between the home and UN conditions, however, reached significance (with the amplitude of the N400 effect being reduced in UN condition) over the frontal electrodes ($F_{frontal} = 8.00$, $p = .014$), but not over the central and parietal electrodes ($F_{central} = 2.29$, $p = .195$; $F_{parietal} = 0.57$, $p = .451$).

We further compared the relative onset latency of the N400 effects across the three dialects, especially between the FA and UN regional dialects. This is because the two regional dialect conditions elicited a reduced negativity only within the N400 window latency. The dependent variable Onset Latency of the N400 effect was assessed by determining the time point where 50% of the peak amplitude was exceeded in the difference waveform (strong-constraint vs. weak-constraint) within 220–500 ms post-noun onset, accompanied with Jackknife correction (Miller, Patterson, & Ulrich, 1998). Analysis of covariance (ANOVA) was conducted with Dialect × Region (central "averaged across *C5/C6, C3C4, C2/C1, CZ, CP5/CP6, CP3/CP4, CP2/CP1, CPZ*" vs. parietal "averaged across *P7/P8, P5/P6, P3/P4, P1/P2, PZ, PO7/PO8, PO5/PO6, PO3/PO4, POZ*") as independent factors. The central and parietal electrodes were selected because the Semantic Constraint effect reached significance in the central-parietal scalp regions in all of the three dialect conditions. This analysis showed a significant interaction between Dialect and Region ($F(2,46) = 17.06$, $p < .001$, $\eta 2 = .426$) due to the fact that the effect of Dialect reached significance only over the central electrodes ($F_{central}(2,46) = 26.42$, $p < .001$, $\eta 2 = .535$; $F_{parietal}(2,46) = .69$, $p = .507$, $\eta 2 = .029$). Further pairwise comparisons over the central electrodes showed that the onset latency of the N400 effect was significantly delayed in the UN regional dialect condition (320 ms [SD = 61]), compared to the FA regional dialect condition ($p < .005$; 249 ms [SD = 37]) and the home dialect condition ($p < .001$; 238 ms [SD = 12]). There was, however, no significant difference between the home and FA regional dialect conditions ($p = .277$).
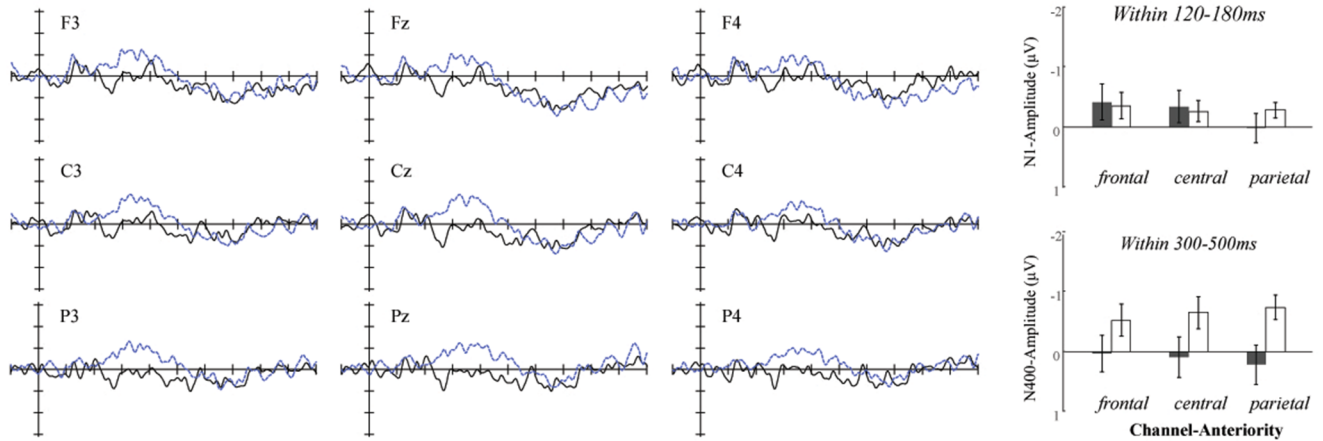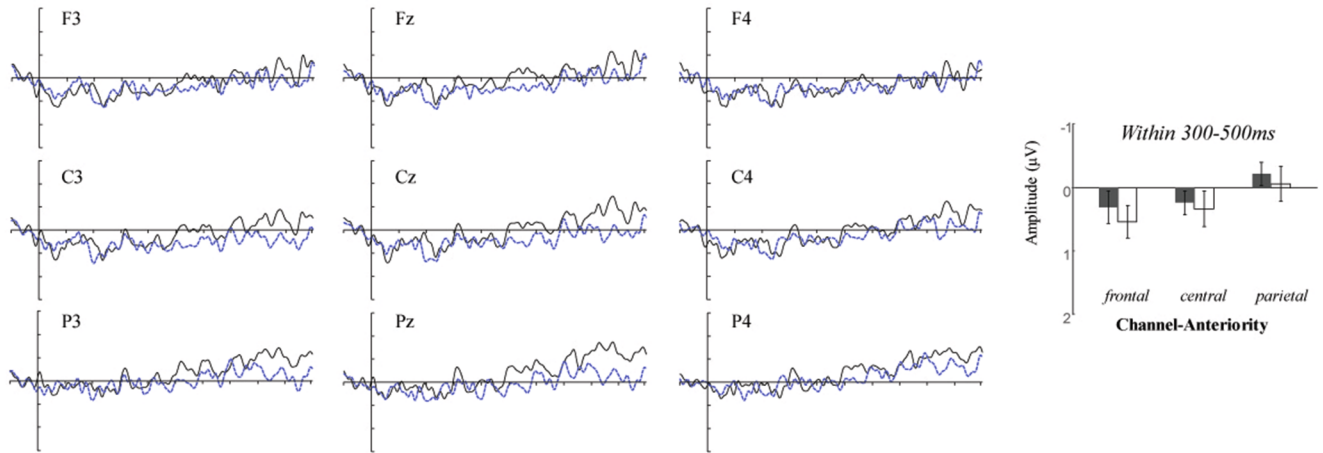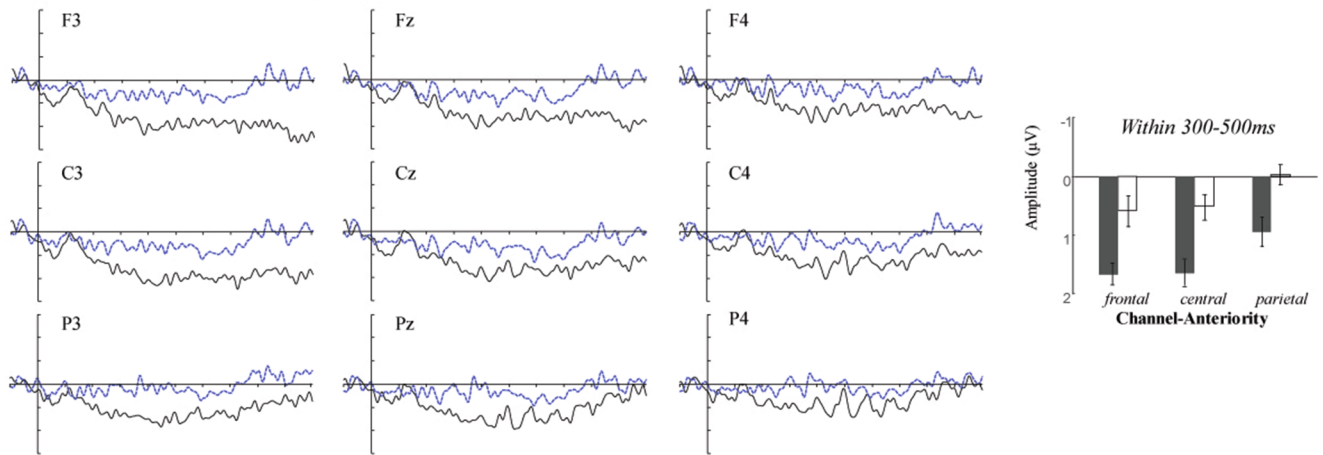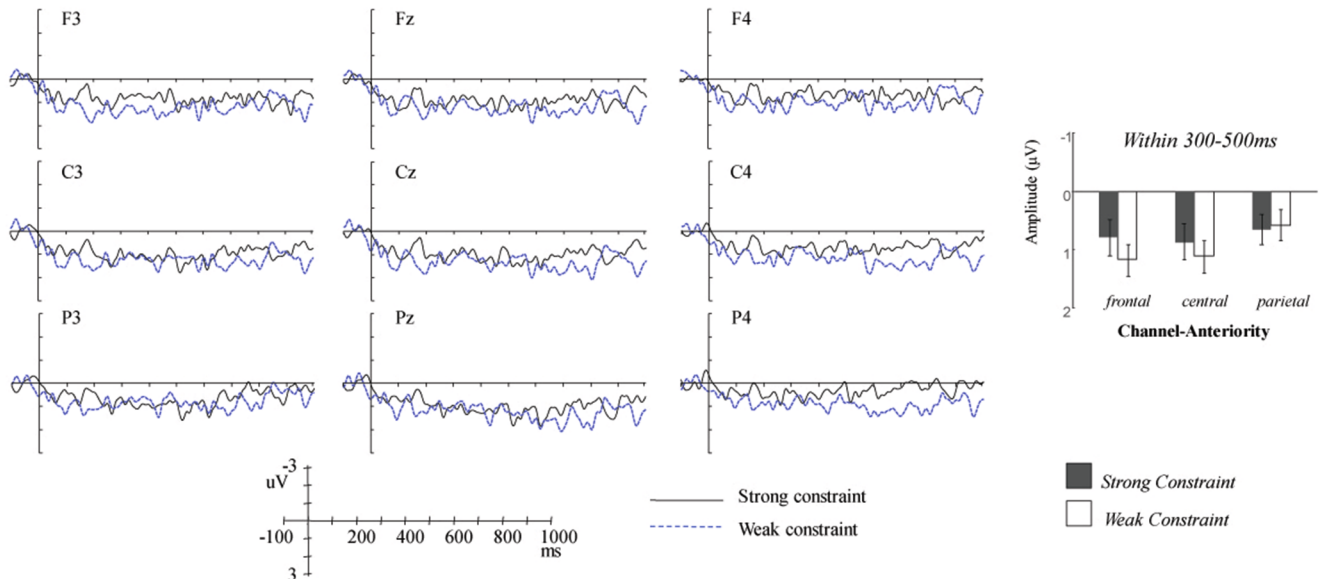
Fig. 2. Grand-average ERPs time-locked to critical nouns in Experiment 1, which were evoked in the case of familiar regional dialect (Beijing Mandarin) (A), home dialect (Dalian Mandarin) (B), and unfamiliar regional dialect (Jiaozuo Mandarin) (C), respectively. The bar plots at the right side indicate the average ERP amplitude (with standard error across participants) within the window latency of 120–180 ms (or 300–500 ms).

ERP results at the critical verbs

Cluster-based random permutation test results

The cluster-based random permutation test revealed a significant

interaction between Semantic Constraint and Dialect, as indicated by the significant difference in the difference waveforms (strong-minus-weak) between the home and FA regional dialects (p < .01, for a time-window of approximately 360 ms to 580 ms post-verb onset) and

A Familiar Regional dialect (Beijing Mandarin)

B Home dialect (Dalian Mandarin)

C Unfamiliar Regional dialect (Jiaozuo Mandarin)

Fig. 3. Grand-average ERPs time-locked to critical verbs in Experiment 1, which were evoked in the case of familiar regional dialect (Beijing Mandarin) (A), home dialect (Dalian Mandarin) (B), and unfamiliar regional dialect (Jiaozuo Mandarin) (C), respectively. The bar plots at the right side indicate the average ERP amplitude (with standard error across participants) within the window latency of 300–500 ms.

between the home and UN regional dialects ($p = .042$, for a time window of approximately 360 ms to 460 ms post-verb onset). The follow-up simple effect analysis showed that, in the case of home dialect

(Dalian Mandarin), there was a significant difference between the strong- and weak-constraint conditions ($p < .005$, at an interval of approximately 260 ms to 496 ms post-verb onset); this effect was most

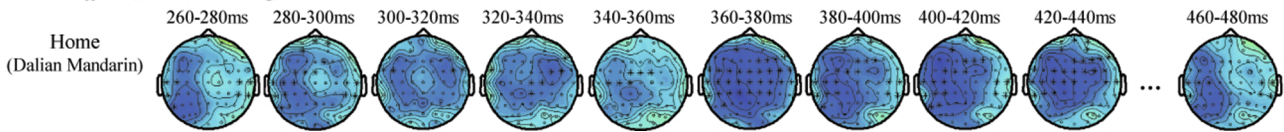**Experiment 1**

A: *ERP effect (Weak vs. Strong) at nouns*
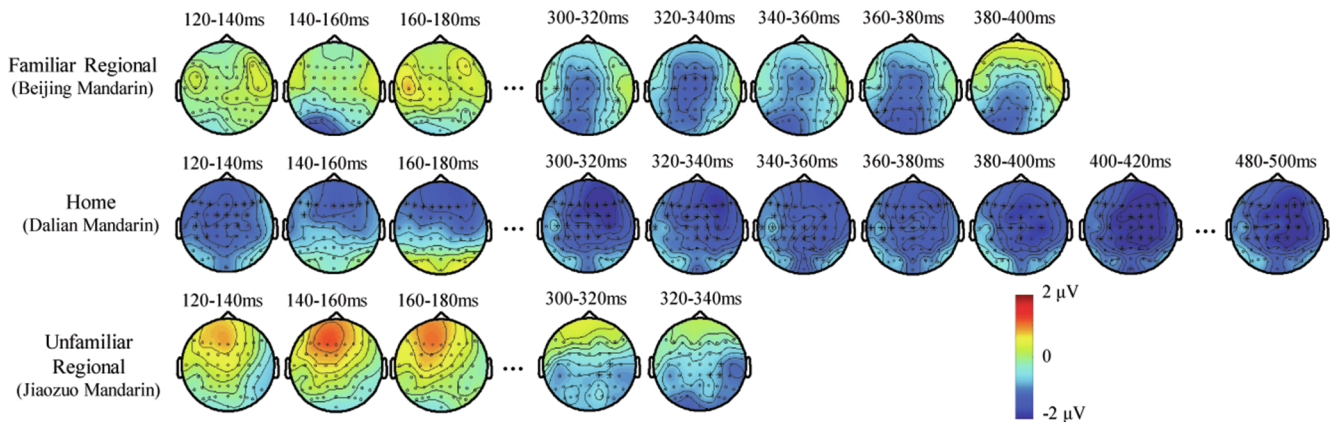
B: *ERP effect (Weak vs. Strong) at verbs*

Fig. 4. Results of the cluster-based permutation test in Experiment 1. Topographies and colors reflect the spatial distribution and magnitude of Semantic constraint effects (weak-minus-strong). The dots indicate the electrodes that were included in the analysis, and the stars indicate the electrodes that showed a significant effect of Semantic Constraint. A: ERP effect (weak-minus-strong) at the nouns in the three dialect conditions; B: ERP effect (weak-minus-strong) at the verbs in the home dialect condition.

pronounced over the frontal-central electrodes. In contrast, in listening to the two regional dialects, we did not find any differences in ERP amplitude between the strong- and weak-constraint conditions (all $ps > 0.104$) (see Figs. 3 and 4).

Linear mixed-effects model analysis results

Like the analysis at the nouns, at the verbs, linear mixed-effects model analyses were performed separately for the frontal, central, and parietal electrode regions. At the verbs, the linear mixed-effects model analysis was conducted only in the 300–500 ms window latency, as the early window latency showed no significant semantic constraint effect in the cluster-based random permutation test analysis.

Within the 300–500 ms window latency, firstly, the categorical analysis found a significant interaction between Semantic Constraint and Dialect in all three scalp regions ($F_{frontal} = 4.61$, $p = .015$; $F_{central} = 4.58$, $p = .015$; $F_{parietal} = 3.37$, $p = .035$). Further simple effect analysis found that, compared to the weak-constraint condition, the strong-constraint condition elicited a significantly enhanced positive deflection in the case of the home dialect ($F_{frontal} = 5.28$, $p = .022$; $F_{central} = 8.12$, $p < .01$; $F_{parietal} = 7.63$, $p < .01$) but not for the two regional dialects (all $ps > .253$). In other words, the mixed-effect analysis revealed result patterns similar to those in the cluster-based permutation test. Secondly, the continuous analysis of single-trial data (linear mixed-effects regression analysis) revealed that in listening to the home dialect, larger positive deflections in the 300–500 ms window were significantly associated with a larger value of cloze2 (frontal: $\beta = 2.33$, CI [0.80, 3.85], $F = 8.93$, $p < .01$; central: $\beta = 2.02$, CI [0.59, 3.46], $F = 7.64$, $p < .01$; parietal: $\beta = 1.66$, CI [0.406, 2.93], $F = 6.67$, $p < .01$). Additionally, a larger ERP amplitude was significantly associated with a larger value of verb cloze probability over the central electrodes ($\beta = 2.06$, CI [0.58, 3.54], $F = 7.43$, $p = .025$). In contrast, neither the value of cloze1 nor the value of verb congruency showed a significant correlation with the averaged EEG voltage in the 300–500 ms window (all $ps > .159$).

In summary, results in Experiment 1 showed that at the critical verbs, the strong-constraint condition (relative to the weak-constraint

condition) elicited an enhanced positive deflection only when listening to the home dialect (Dalian Mandarin). The amplitude increases of this positivity were significantly correlated with increases in the predictability of forthcoming critical nouns. This correlation effect reached the maximum at the frontal electrodes as indicated by the linear mixed-effects regression analysis. At the critical nouns, although the highly predictable nouns (compared to the weakly predictable nouns) elicited a reduced N400 for all three dialects (home, FA, and UN), N1 was reduced only for the home dialect. Given that the enhanced positivity at the verbs and the reduced N1 at the nouns were observed only in the case of the home dialect, one may argue, as one anonymous reviewer pointed out, that these effects could be merely driven by the low-level acoustic differences such as the intrinsic acoustic-phonological properties of the critical verbs and nouns in the home Dalian Mandarin condition, hence indicating no cognitive significance. Although the linear mixed effects modelling of the single trials in our ERP analyses have taken care of the potential effects of stimulus acoustic properties on ERP responses, to rule out the acoustic-difference explanation for the ERP effects observed in the home dialect (Dalian Mandarin) condition in Experiment 1, we ran the same experiment (Experiment 2) with participants whose home dialect is Beijing Mandarin.

Experiment 2

All participants in Experiment 2 were born and grew up in the city of Beijing. Like the Dalian Mandarin speakers in Experiment 1, the Beijing Mandarin listeners listened to and comprehended the same set of stimuli produced in Beijing Mandarin, Dalian Mandarin, or Jiaozuo Mandarin. These experimental stimuli also varied in terms of the semantic constraint in sentence context (strong vs. weak). However, in Experiment 2, Beijing Mandarin was the home dialect. Although Dalian Mandarin shares some more phonetic/phonological similarities to Beijing Mandarin than Jiazuo Mandarin, both were treated as unfamiliar regional dialects (UN regional dialects), as both are at best only rarely encountered by the Beijing Mandarin listeners.

If the enhanced positivity and reduced N1 observed in the home

dialect (Dalian Mandarin) condition in Experiment 1 were merely driven by the stimuli's acoustic properties, these effects would again be observed only in the case of Dalian Mandarin in Experiment 2. In contrast, if these ERP effects were associated with a cognitive significance driven by home-dialect experience, we would expect that in Experiment 2, similar effects should be observed when participants listen to their home dialect (i.e., Beijing Mandarin). Furthermore, effects observed only in Dalian Mandarin (i.e. the home dialect condition) in Experiment 1 should not be observed for Beijing Mandarin participants in Experiment 2.

Results

Behavioral results

For the comprehension questions in each dialect, each participant included in the final statistical analysis of the EEG data showed an accuracy rate higher than 75% (averaged accuracy rate: 94.21% [SD = 4.78], 90.04% [SD = 7.68], 84.72% [SD = 7.00] for Beijing Mandarin [home dialect], Dalian Mandarin, and Jiaozuo Mandarin, respectively). The ERP results are reported below and summarized in Table 5.

Table 5
Summary of the ERP effects at the critical transitive verbs and target object nouns in Experiment 2.

	Time window	Description of significant effects	Polarity	Topography
Verbs	100–160 ms	No significant effect	N.A.	N.A.
	300–500 ms	*Significant interaction (S*D):* Enhanced negative deflection for strong-constraint in home dialect only (associated with cloze2)	Negativity	Frontal-central distribution, with maximum at frontal electrodes
Nouns	120–180 ms	*Significant interaction (S*D):* Reduced N1 for strong-constraint in home dialect only (associated with cloze3)	Negativity	Central-parietal distribution
	300–500 ms	*Significant main effect (S):* Reduced N400 for strong-constraint in all of the three dialects (associated with cloze3)	Negativity	Central-parietal electrodes

Note. S and D indicate the 'Semantic Constraint' and ''Dialect', respectively.

Material and methods

Ethics statement

This research was approved by the Ethics Committee of Institute of Psychology, Chinese Academy of Sciences. All participants were over 18 years of age and gave written informed consent. They were notified that their participation was completely voluntary and that they could secede at any time.

Participants

Thirty-one right-handed university students (ten males), all of whom were born and grew up in the city of Beijing (China), participated in this experiment. Their mean age was 20 years (range 19–28). None reported any medical, neurological, or psychiatric illness, and all gave informed consent. They were all native speakers of Beijing Mandarin. The data of seven participants (three males) were removed from analysis because of excessive artifacts (five participants) or low accuracy of behavioral responses (two participants). Therefore, data of twenty-four subjects were included in the final statistical analysis.

Materials, procedure, and EEG acquisition

Materials and EEG acquisition procedures were the same as those in Experiment 1.

ERP and statistical analyses

The methods for ERP and statistical analyses were the same as those in Experiment 1. Trials containing the artifacts mentioned above were rejected. The ANOVA test with conditions (six levels) as independent factors showed that there was no significant difference in the rejected trials of the six experimental conditions for both nouns ($F_{(5,115)}$ = 1.25, p = .291) (8.8% [SD = 5.6], 9.1% [SD = 6.6], 8.6% [SD = 6.0], 7.3% [SD = 5.7], 8.8% [SD = 7.0], and 6.5% [SD = 5.2] for Beijing-Strong, Beijing-Weak, Dalian-Strong, Dalian-Weak, Jiaozuo-Strong, and Jiaozuo-Weak conditions, respectively) and verbs ($F_{(5,115)}$ = 0.69, p = .575) (8.9% [SD = 6.8], 9.8% [SD = 5.3], 8.8% [SD = 6.0], 9.1% [SD = 5.5], 9.8% [SD = 6.7], and 7.9% [SD = 7.3] for Beijing-Strong, Beijing-Weak, Dalian-Strong, Dalian-Weak, Jiaozuo-Strong, and Jiaozuo-Weak conditions, respectively).

ERP results at the critical nouns

Cluster-based random permutation test results

For the comparison of difference waveforms (strong-minus-weak) between the home dialect and the regional Dalian dialect, a cluster in the data was found to extend from 140 ms to 210 ms post-noun onset and another cluster from 230 ms to 340 ms; the cluster-based random permutation test indicated that there was a significant difference between the home and regional Dalian dialect conditions (p = .012). For the comparison of strong-minus-weak difference waveforms between the home and regional Jiaozuo dialect conditions, one cluster was found to extend from 160 ms to 210 ms and another cluster from 230 ms to 390 ms; the cluster-based random permutation test demonstrated that the difference between the home and regional Jiaozuo dialect conditions reached significance (p = .033). These results indicated that there was a significant interaction between Semantic Constraint and Dialect.

The subsequent simple effect analysis demonstrated that when participants listened to the home Beijing dialect, we observed a significant difference between the strong- and weak-constraint conditions (p < .001, for a time window of approximately 118 ms to 480 ms post-noun onset); this effect was most pronounced over the central-parietal electrodes. For the two regional dialects, there was also a significant difference between strong- and weak-constraint conditions (Dalian Mandarin: p = .041, at an interval of approximately 320 ms to 410 ms post-noun onset; Jiaozuo Mandarin: p = .046, at an interval of approximately 360 ms to 420 ms post-noun onset); these effects were most pronounced over the central-parietal electrodes (see Figs. 5 and 7).

In the case of the home Beijing dialect, the ERP deflections underlying the sustained semantic constraint effect had two clearly separate peaks, one in the N1 window latency and the other in the N400 widow latency (see Fig. 5). Therefore, we classified this ERP effect as a combination of the N1 and N400 effects. In contrast, the effects in the two regional dialect conditions were observed within the window latency of N400 and hence classified as the N400 effect.

Linear mixed-effects model analysis results

Based on the results of the cluster-based random permutation tests, linear mixed-effects analyses were conducted separately for the frontal,
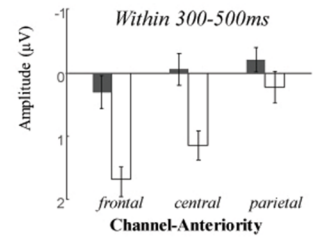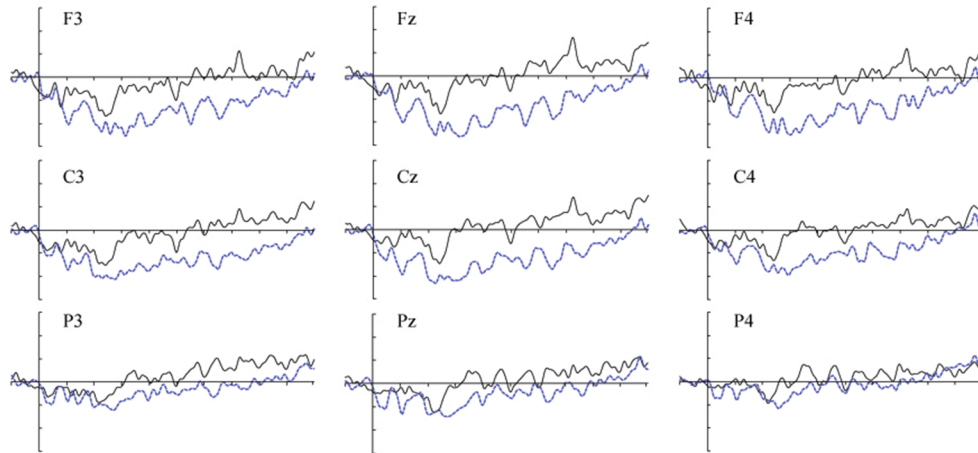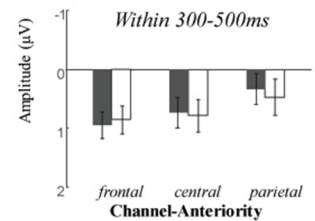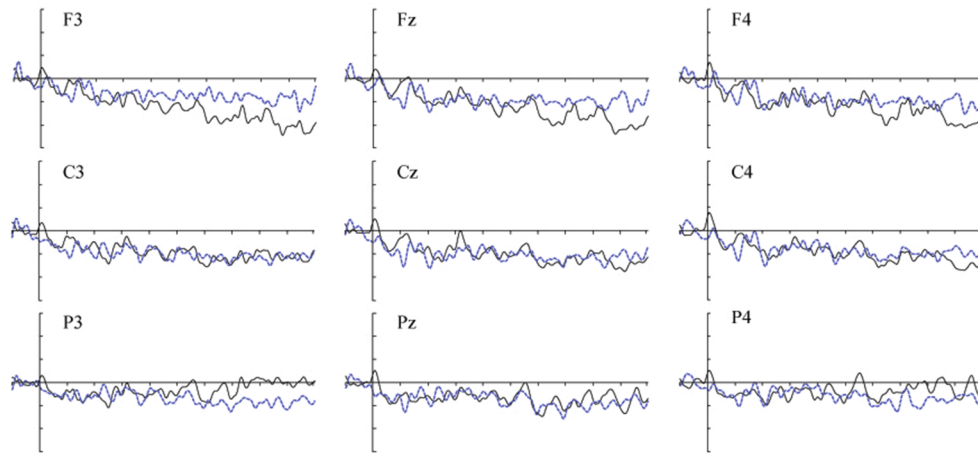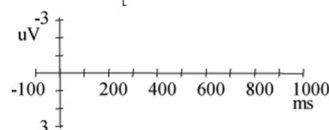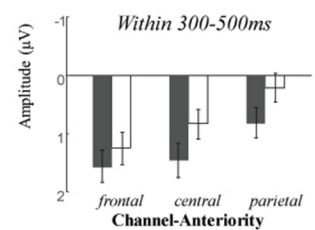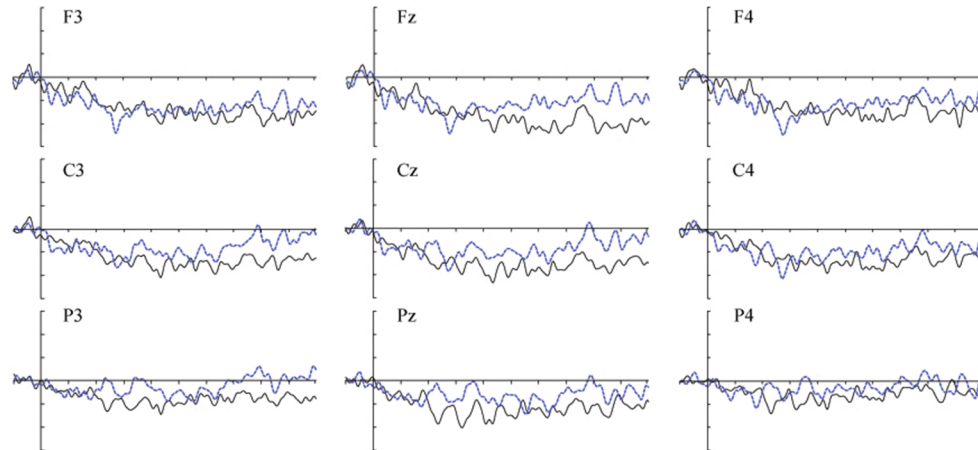
A Home dialect (Beijing Mandarin)
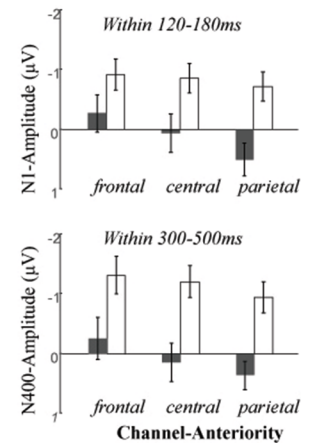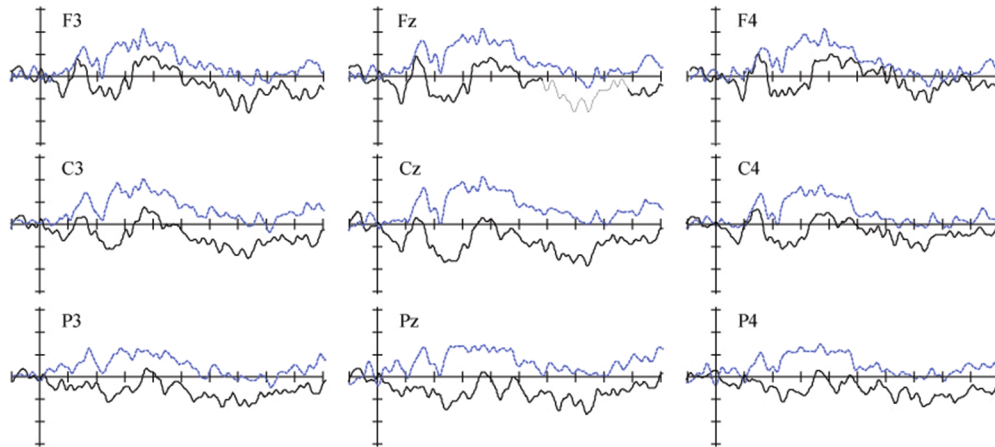
B Unfamiliar regional dialect (Dalian Mandarin)
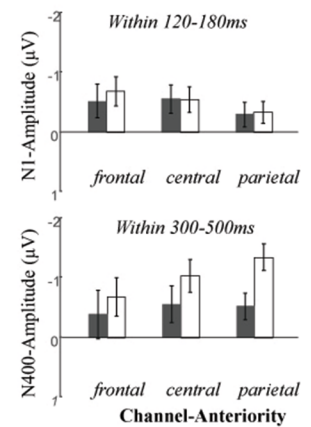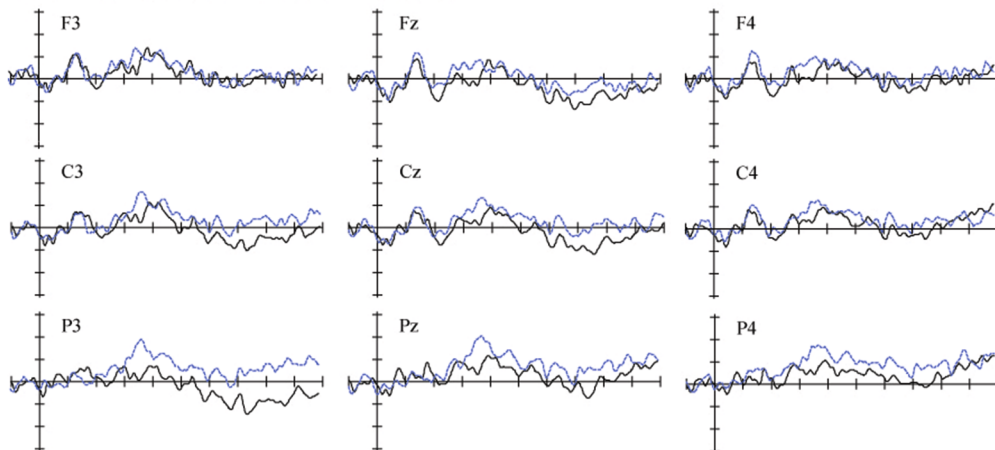
C Unfamiliar Regional dialect (Jiaozuo Mandarin)
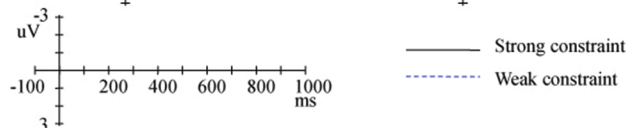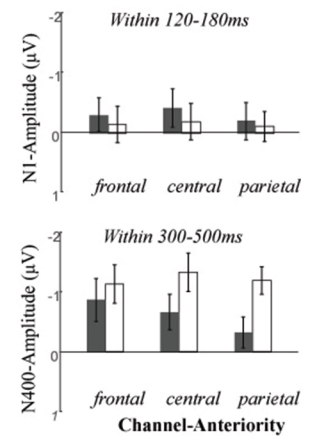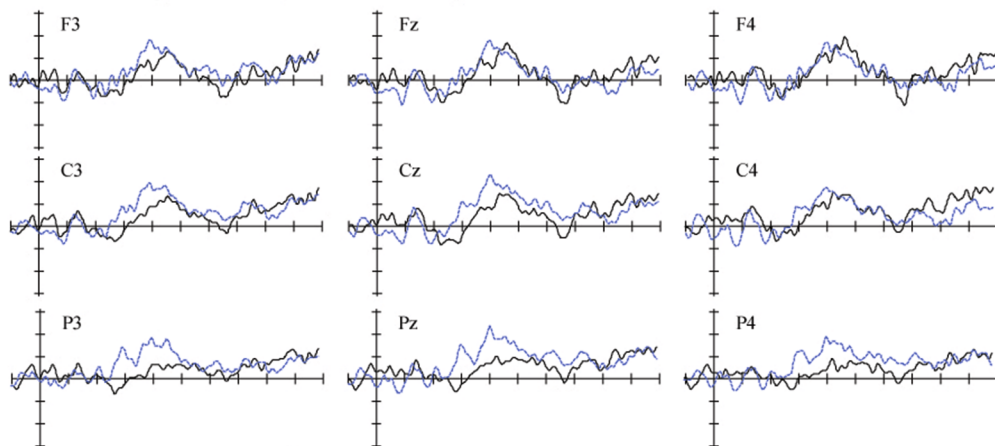
Fig. 5. Grand-average ERPs time-locked to critical nouns in Experiment 2, which were evoked in the case of home dialect (Beijing Mandarin) (A), unfamiliar regional dialect (Dalian Mandarin) (B), and unfamiliar regional dialect (Jiaozuo Mandarin) (C), respectively. The bar plots at the right side indicate the average ERP amplitude (with standard error across participants) within the window latency of 120–180 ms (or 300–500 ms).

central, and parietal scalp regions, as well as for the N1 (120–180 ms) and N400 (300–500 ms) time windows.

Within the N1 window latency, firstly, categorical analyses revealed

a significant interaction between Semantic Constraint and Dialect in the central and parietal regions ($F_{central} = 3.98$, $p = .028$; $F_{parietal} = 5.93$, $p < .01$). In the frontal regions, neither the main effect of Semantic

Fig. 6. Grand-average ERPs time-locked to critical verbs in Experiment 2, which were evoked in the case of home dialect (Beijing Mandarin) (A), unfamiliar regional dialect (Dalian Mandarin) (B), and unfamiliar regional dialect (Jiaozuo Mandarin) (C), respectively. The bar plots at the right side indicate the averaged ERP amplitude (with standard error across participants) within the window latency of 300–500 ms.

Constraint nor its interaction with Dialect reached significance (all $ps > .187$). Further simple effect analysis demonstrated that, compared to the weak-constraint condition, the strong-constraint condition

elicited a significantly reduced N1 for the home dialect ($F_{central} = 7.75$, $p < .01$; $F_{parietal} = 14.59$, $p < .001$) but not for the two regional dialect conditions (all $ps > .845$). Further continuous analysis of
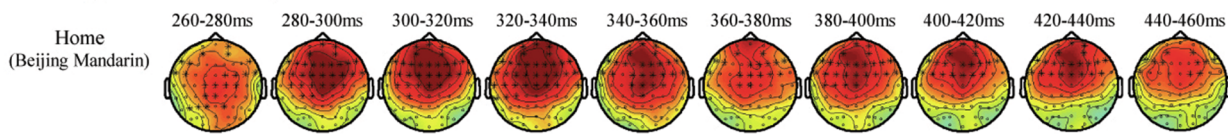
Experiment 2

A: *ERP effect (Weak vs. Strong) at nouns*
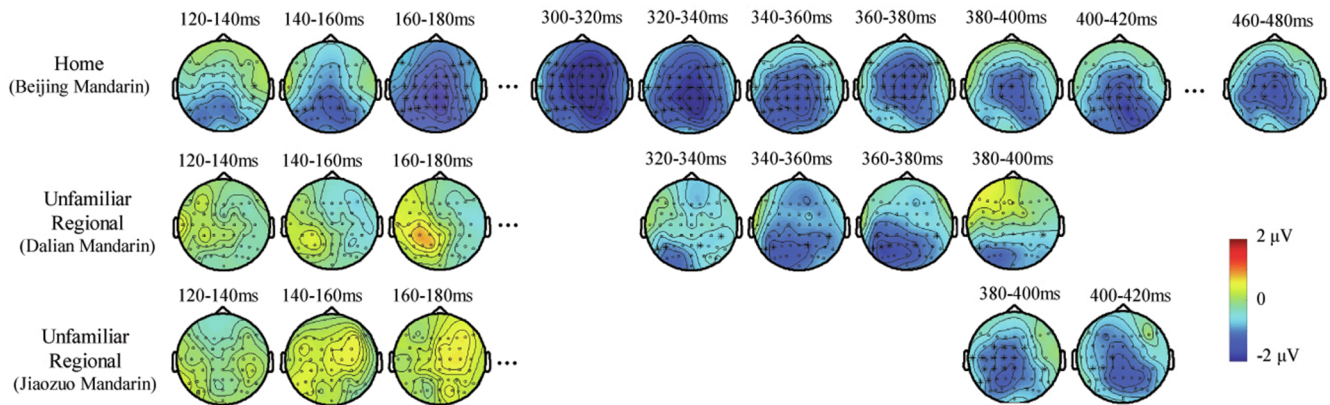
B: *ERP effect (Weak vs. Strong) at verbs*

Fig. 7. Results of the cluster-based permutation test in Experiment 2. Topographies and colors reflect the spatial distribution and magnitude of Semantic constraint effects (weak-minus-strong). The dots indicate the electrodes that were included in the analysis, and the stars indicate the electrodes that showed a significant effect of Semantic Constraint. A: ERP effect (weak-minus-strong) at the nouns in the three dialect conditions; B: ERP effect (weak-minus-strong) at the verbs in the home dialect condition.

single-trial data found that when listening to the home dialect, higher cloze3 values were strongly associated with smaller negative deflections (N1) (central: $\beta = 1.87$, CI [0.24, 3.49], $F = 5.06$, $p = .032$; parietal: $\beta = 2.14$, CI [0.69, 3.58], $F = 8.41$, $p = .015$), but neither the noun congruency (all $ps > .317$) nor the other factors' effects (all $ps > .544$) reached significance.

Within the N400 window latency, firstly, the categorical analyses resulted in a significant main effect of Semantic Constraint over the central and parietal electrodes ($F_{central} = 10.19$, $p < .01$; $F_{parietal} = 17.14$, $p < .001$; $F_{parietal} = 3.30$, $p = .082$), indicating that the strong-constraint condition elicited a reduced N400 (compared to the weak-constraint condition). The Semantic Constraint × Dialect interaction did not reach significance ($F_{frontal} = 1.22$, $p = .443$; $F_{central} = 1.34$, $p = .443$; $F_{parietal} = 0.398$, $p = .672$), suggesting that the magnitude of the semantic-constraint-related N400 effect was not modulated by different dialects. Secondly, the continuous analyses showed that a smaller N400 deflection was significantly associated with larger values of cloze3 (central: $\beta = 1.77$, CI [0.90, 2.63], $F = 16.01$, $p < .001$; parietal: $\beta = 1.72$, CI [0.87, 2.58], $F = 15.56$, $p < .001$) and a larger stroke number of nouns (central: $\beta = 6.62$, CI [1.76, 12.32], $F = 6.80$, $p < .01$; parietal: $\beta = 7.54$, CI [3.03, 12.57], $F = 9.57$, $p < .005$); however, neither the effect of noun congruency (all $ps > .485$) nor the other factors reached significance (all $ps > .152$).

ERP results at the critical verbs

Cluster-based random permutation test results

The cluster-based random permutation test revealed a significant interaction between Semantic Constraint and Dialect, as indicated by the significant differences in the different waveforms (strong-minus-weak) between the home Beijing Dialect and the regional Dalian dialect ($p = .036$, at an interval of approximately 280 ms to 360 ms post-verb onset) and between the home and Jiaozuo dialect ($p < .01$, at an interval of approximately 290 ms to 500 ms post-verb onset). The follow-up simple effect analysis showed that in the case of home dialect

(Beijing Mandarin), the difference between the strong- and weak-constraint conditions reached significance ($p < .001$, at the interval of approximately 260 ms to 480 ms); this effect was most pronounced over the frontal-central electrodes. In contrast, when participants listened to the two regional dialects, we observed no significant difference in the ERP amplitude between the strong- and weak-constraint conditions (all $ps > 0.105$) (see Figs. 6 and 7).

Results of linear mixed-effects model analysis

Linear mixed-effects model analyses were conducted separately for the frontal, central, and parietal scalp regions within the 300–500 ms window latencies.

Within the 300–500 ms window latency, firstly, the categorical analysis showed a significant interaction between Semantic Constraint and Dialect in the frontal and central scalp regions ($F_{frontal} = 5.88$, $p < .01$; $F_{central} = 5.43$, $p < .01$). At the parietal channels, neither the main effect of Semantic Constraint ($F_{parietal} = 0.008$, $p = .929$) nor its interaction with Dialect reached significance ($F_{parietal} = 2.17$, $p = .114$). Further simple effect analysis in the frontal and central scalp regions found that, compared to the weak-constraint condition, the strong-constraint condition elicited a significantly enhanced negative deflection in the home dialect condition ($F_{frontal} = 11.52$, $p < .005$; $F_{central} = 8.00$, $p < .005$) but not in the two regional dialects (all $ps > .290$). The categorical analysis of single-trial data revealed similar results to those of the cluster-based permutation test. Secondly, further continuous analysis in the frontal and central scalp regions revealed that when listening to the home dialect, larger cloze2 values were strongly associated with larger negative deflections in the 300–500 ms window ($\beta = -2.38$, CI [-3.94, -0.82], $F_{frontal} = 8.97$, $p < .01$; $\beta = -1.81$, CI [-3.29, -0.33], $F_{central} = 5.78$, $p = .016$); in contrast, neither the values of cloze1 nor the values of verb congruency (or verb cloze) were significantly correlated with the averaged EEG voltages in the 300–500 ms window (all $ps > .227$).

In summary, the results of Experiment 2 showed that at the critical verbs, the strong-constraint condition (compared to the weak-constraint condition) evoked an enhanced anterior negativity (around the N400

window latency) only when listening to the home dialect (Beijing Mandarin), which significantly correlated with the predictability of the incoming nouns. At the critical nouns, the strongly predictable nouns (compared to the weakly predictable ones) elicited a reduced N400 for all three dialect conditions but an early N1 reduction only in the home dialect (Beijing Mandarin) condition. These results indicate the cognitive significance of our findings with regard to the role of dialectal experience in predictive processing and rule out the possibility that the effects observed in Experiment 1 (as well as in Experiment 2) were merely driven by low-level acoustic differences.

Note that the polarity of the ERP components (or ERP effects) evoked at the critical verbs in Experiments 1 and 2 was not exactly the same. The strong-constraint condition evoked an enhanced positive ERP response in Experiment 1 but an enhanced N400-like response in Experiment 2. Despite the polarity differences across the two experiments, both peaked over frontal electrodes and were observed within a similar time window (i.e., 300–500 ms post-verb onset). Furthermore, both correlated significantly with the cloze probability of the incoming critical nouns. Conjointly, these ERP effects at the verb suggest a cognitive significance about the anticipatory processing of the following object nouns, which we will discuss in more detail in the following section.

General discussion

The present EEG study investigated how dialectal experience modulates anticipatory lexical processing and how this knowledge sheds light on the general mechanisms of predictive speech processing. Two experiments were conducted with the same set of stimuli, differing only in the listeners' home dialect. We not only examined the processing of more predictable vs. less predictable target nouns but also that of the transitive verbs that precede the target object nouns. At the pre-nominal transitive verbs, both experiments demonstrated that only when listening to the home dialect did the strong-constraint condition (compared to the weak-constraint condition) elicited significant frontal-central ERP responses, reflected as an enhanced positivity in Experiment 1 but as an enhanced N400-like response in Experiment 2. At the target object nouns, both experiments showed that, in addition to the N400 reduction effect that appeared in all three dialect conditions, the strongly predictable object nouns also led to a negativity reduction in the early N1 window latency albeit only in the home dialect condition (i.e., Dalian Mandarin in Experiment 1 and Beijing Mandarin in Experiment 2). These results and their implications for models of lexical predictive processing are discussed below.

The facilitating effect of semantic context on the processing of newly available words

The consistent reduction in N400 amplitude at the critical nouns in the strong-constraint condition for all three dialects suggests that the semantic processing of strongly predictable nouns was facilitated regardless of dialect experience. Moreover, the N400 amplitude was significantly correlated with the predictability of these nouns before they appeared in the context, which provides direct evidence that the facilitated processing of the target nouns was driven by the preceding semantic constraint. Note that the nouns' semantic congruency in the present study was not found to significantly modulate the N400 amplitude at these nouns, which might be because all critical nouns in our study had a relatively high degree of semantic congruency and small variance. Nevertheless, our results are consistent with previous studies that manipulated the cloze probability of words and showed that a strongly constraining semantic context can facilitate the later semantic processing of incoming new words (e.g., Besson, Kutas, & Petten, 1992; Diaz & Swaab, 2007; Federmeier, 2007; Laszlo & Federmeier, 2009; Szewczyk & Schriefers, 2013; Thornhill & Van Petten, 2012).

The N400 effect observed at the nouns in the home dialect of Experiment 1 had a broad scalp distribution, which was more frontally distributed relative to the standard central-parietal N400 effect observed in the other dialect conditions. Though more frontally distributed, this negativity effect was classified as N400 since it peaked around 400 ms, given that previous studies have also reported similar more frontally distributed N400 effects (Balconi & Caldiroli, 2011; Guillaume, Tinard, Baier, & Dufau, 2016; Kelly, Ward, Creigh, & Bartolotti, 2007; Nakano, Saron, & Swaab, 2010). The scalp distribution difference observed across the three dialect conditions in the present study suggests some differences in the neural basis of the N400 effects during lexical prediction, which needs to be examined further.

Importantly, when the message was spoken in the listeners' home dialect (i.e., Dalian Mandarin in Experiment 1 and Beijing Mandarin in Experiment 2), the strongly predictable target nouns also elicited a smaller N1 (compared to the weakly predictable ones). It is well known that the N1 component is a robust auditory component and sensitive to the acoustics of individual stimulus (Näätänen & Picton, 1987), especially the acoustic/phonological perceptual features present in the speech signal such as vowels, syllables, and word onsets (Obleser et al., 2004; Parviainen et al., 2005; Sanders & Neville, 2003).

In this study, the N1 effects were not likely to be caused by the intrinsic acoustic-perceptual properties of the stimulus words in a specific dialect condition due to the following reasons. First, the acoustic cues of the critical nouns (such as duration, intensity, and pitch maximum) were carefully matched between the strong- and weak-constraint conditions for all three dialects. All listeners were exposed to exactly the same set of auditory stimuli, but the reduced N1 effect was observed only when listeners were processing their home dialect (i.e., Dalian Mandarin in Experiment 1 and Beijing Mandarin in Experiment 2). Secondly, the regression analyses showed that decreases in the N1 amplitude at the target nouns correlated significantly with increases in the cloze probability of these nouns after controlling for the potential influence of acoustic cues.

It is also unlikely that the reduced N1 effects were mere spill-over effects from pre-target ERP effects. In Experiment 1, the ERP effects time-locked to the verbs ended before 600 ms after the transitive verb onset, i.e., before the following target nouns appeared (as the duration interval between the onset of the verbs and that of the nouns was on average about 800 ms). Note that the duration interval between the onset of critical verbs and that of critical nouns in both experiments varied between trials, which functioned as a time jitter and, consequently, reduced/cancelled out the baseline differences if existed. Furthermore, for ERPs at the target nouns, baseline correction was performed by subtracting the ERP variance within 100 ms preceding the onset of the target nouns. If the N1 effects at the target nouns were due to the carryover of processing earlier linguistic input, we should have observed a reversed effect at the nouns (compared to that at the verbs), which is inconsistent with the present results. In Experiment 2, the ERP effects at the verbs had a frontal-central scalp distribution, while the ERP effects at the nouns had a central-parietal distribution (reaching maximum at the parietal channels), again rendering it unlikely that there was a spill-over effect from the processing of the verbs to that of the nouns.

Therefore, we feel it safe to conclude that the N1 reduction effect indicates that in the home dialect condition, a strong-constraint semantic context facilitates the early acoustic/perceptual processing (pre-semantic processing) of the incoming new words. This early facilitating effect is in line with the findings reported in Sohoglu et al. (2012), which demonstrated that the facilitating effect of prior knowledge from matching text (e.g., *clay-clay* for matching, *snail-glass* for mismatching, and *xxx-haze* for neutral) is already present in the early 90–130 ms time window of the following spoken word. The early facilitating effect in our study is also consistent with findings in reading comprehension research that show that sentence/discourse context can facilitate the early processing of visual, orthographic, and grammatical features of upcoming words (e.g., Dikker, Rabagliati, & Pylkkänen, 2009; Federmeier & Kutas, 2001; Kim & Lai, 2012; Lee et al., 2012; Molinaro,

Conrad, Barber, & Carreiras, 2010; Penolazzi et al., 2007). Findings in the present study thus extend the early facilitating effect of semantic context constraint on perceptual processing to auditory speech comprehension at the sentence level.

By comparing the facilitating effect of the semantic context on the processing of incoming new words in three different dialects, our results also confirmed the effect of early language exposure (to home dialect) on the semantic-context benefit and how such a benefit is related more generally to anticipatory processing during speech comprehension. In Experiment 1, even though both the home and FA regional dialects were familiar to the listeners, it was only when listening to the home dialect that the facilitating effect of semantic constraint showed an effect in the early N1 window latency. Moreover, although the facilitating effect of semantic context was observed within the N400 window latencies in all three dialects, the onset latency of this N400 reduction effect was delayed in the UN regional dialect, compared to that in the FA and home dialect conditions. This suggests that dialectal familiarity and age of acquisition conjointly affect language processing proficiency and then modulate the semantic context benefit effect, consistent with findings in behavioral studies mentioned earlier (Clopper, 2012; Labov & Ash, 1997). Our results also echo recent ERP studies on second language processing that showed reduced or delayed ERP N400 effects in non-native speech processing (e.g. Martin et al., 2013).

The predictive coding theory (e.g., Friston, 2005, 2010) and related perspectives on language comprehension (e.g., Kuperberg & Jaeger, 2016) have proposed that high-level conceptual/semantic prediction can be quickly transmitted to lower levels of representation and, consequently, can influence early sensory-perceptual processing of new inputs. Our results showed that it was only in the home dialect listening condition that the semantic context benefit at the target nouns was revealed in the early N1 window latency (that is associated with acoustic/perceptual processing). Meanwhile, it was only when our participants listened to their home dialect that anticipatory lexical processing was efficiently implemented at the prenominal verbs (see further discussion in the next section). These results suggest that it is the joint force of lexical-semantic pre-activation and efficient information flow between semantic and acoustic/phonological representations that has led to the early N1 auditory facilitation effects in the early-exposed home dialect.

Anticipatory language processing and the underlying cognitive mechanisms

An important finding of our study is that before the arrival of the critical nouns in a continuous speech signal, significant ERP effects (as a function of the strong vs. weak constraint conditions) can already be observed at the preceding transitive verbs (enhanced positivity in Experiment 1 and enhanced negativity in Experiment 2). These ERP effects could not be accounted for as merely due to the easiness of integrating these verbs into their preceding contexts or due to the low-level acoustic characteristics of the dialectal stimuli. This is because the written version of the corresponding words (namely, the critical verbs and classifiers/adjectives preceding the critical nouns) was exactly the same in the strong- and weak-constraint conditions, and the potential contribution of their acoustic parameters in the spoken data were regressed out by simultaneously modeling them. The fact that the ERP effects were observed only in the case of the native dialect (i.e. Dalian Mandarin in Experiment 1 and Beijing Mandarin in Experiment 2) also argues against the acoustic account.

More importantly, our single-trial mixed-effect regression analysis showed that the ERP responses at the critical verbs (within 300–500 ms post-verb onset) were significantly predicted by the following critical nouns' predictability even after we strictly controlled and statistically regressed out the potential contributions of the verbs' semantic congruency or the constraint of their preceding context. Taken together, a more reasonable interpretation of the ERP effects at the critical verbs is that after hearing the critical verbs, listeners are able to combine the

verb meaning with preceding context (and/or long-term knowledge) to generate hypothesized lexical meanings of the forthcoming nouns.

Note that the polarities of the ERP components evoked at the critical verbs in the two experiments are not exactly the same, with the strong-constraint condition evoking an enhanced positive ERP deflection in Experiment 1 but an enhanced negative ERP deflection in Experiment 2. Despite these polarity differences, both peaked over frontal electrodes and were observed within a similar time window (i.e., around 300–500 ms post-verb onset). More importantly, for both, variations in ERP amplitude significantly correlated with variations in the cloze probability of the following target nouns. This suggests that the ERP effects in the two experiments are likely to be associated with the same cognitive significance in terms of the anticipatory processing of incoming nouns. However, the specific neural mechanisms that support this anticipatory processing might differ in the two experiments. Further research is needed to understand the source(s) and nature(s) of the specific cognitive process(es). In spite of the polarity differences, the ERP effects at the critical verbs suggest that when confronted with the home dialect and a strongly constraining semantic context, the language processing system is able to conduct top-down anticipatory processing to predict and generate the lexical-semantic features of the incoming words in the continuous speech signal.

As mentioned in the introduction, previous studies have investigated predictive lexical processing in sentence/discourse comprehension by measuring ERPs on pre-nominal materials (e.g., articles or gender markers before nouns) (e.g., DeLong et al., 2005; Foucart et al., 2014; Ito et al., 2017; Kochari & Flecken, 2019; Martin et al., 2013; Molinaro et al., 2017; Nieuwland et al., 2018). These studies typically examine lexical pre-activation by measuring ERPs evoked by the article, determiner, or the grammatical form of the following target noun. Unlike these studies, the present study manipulated the semantic context constraint and measured ERPs elicited by the preceding transitive verb of the target noun, which is more directly related to lexical-semantic pre-activation. A similar early effect of lexical prediction (of nouns) on pre-nominal verbs has been reported in previous reading comprehension studies (Li et al., 2017; Szewczyk & Schriefers, 2017). Our results confirmed early findings that the language processing system is able to initiate anticipatory lexical-semantic processing during both visual and auditory mode of processing. The results also extend our understanding of predictive language processing by showing that anticipatory lexical processing in continuous speech comprehension may change dynamically as a function of a specific type of language experience, i.e., a phonetics/phonology-centered dialect experience.

The modulating effect of dialectal experience on anticipatory lexical processing observed in the present study can in turn help us to understand the cognitive mechanism(s) underlying predictive language processing. As mentioned in the introduction, what is still under debate is the extent to which predictions in language comprehension are effortfully generated or just reflect an automatic pre-activation of semantic content from the context (e.g., Baggio, 2018; Huettig, 2015; Kahneman, 2011). The present study showed that direct evidence of anticipatory lexical processing effect was observed only in the home dialect condition, as indicated by both the significant ERP effects at the transitive verbs and the significant correlation between the ERP magnitudes at the verbs and the cloze probability of the following critical nouns. In listening to the two regional dialects, no significant semantic-constraint effect was observed at the transitive verbs.

How to account for the lack of ERP responses at the prenominal verbs in the two regional dialects? One may be tempted to take resort to lexical access difficulties in listening to regional dialects, which impeded pre-nominal anticipatory lexical processing. Such a possibility then makes it irrelevant whether lexical anticipation is based on automatic activation or effortful computation. However, it is important to note that the ERP responses at the target nouns did show a semantic-constraint effect (N400 reduction in the strong- compared to weak-constraint conditions) in all three dialect conditions, raising doubts to

such a general account of lexical accessibility difficulty.

We reckon that lexical access difficulty may indeed be relevant for the UN regional dialect. In this condition, the magnitude of the N400 effect at the nouns was reduced over the frontal electrodes (although no cross-dialect difference was observed over the central-parietal electrodes), as indicated by the results of the categorical single-trial data analysis in Experiment 1. Meanwhile, as revealed by the onset latency analysis in Experiment 1, this N400 effect was also delayed by approximately 70 ms. These results suggest that, in the case of listening to the UN regional dialect, our participants could generally understand the meaning conveyed by the speech signal but their understanding was not as fully as in the home dialect condition. We suspect that there might be tokens that they had lexical access difficulties, contributing to the lack of ERP effects at the pre-nominal verbs.

However, lexical access difficulty is unlikely to be the case in the FA condition (i.e. Beijing Mandarin in Experiment 1). This is because the FA regional dialect used in this experiment is essentially the same as Standard Mandarin Chinese, which is expected to be fully understood (see the introduction section for detailed reasons). Compared to the home dialect condition, the FA condition elicited a comparable N400 reduction effect at the nouns (as indicated by the categorical single-trial data analysis in Experiment 1). This confirms that our listeners were able to successfully access the lexical meanings and comprehend the stimulus sentences. Moreover, the semantic-constraint-related N400 effect at the nouns was observed in the standard N400 window latency (300–500 ms post-noun onset) and the onset latency of this N400 effect (defined by using the Jackknife-based method; Miller et al., 1998) showed no significant difference between the home dialect (starting around 238 ms post-noun onset) and the FA regional dialect (starting around 249 ms). This suggests that while listening to the FA regional dialect, our participants were able to perform the semantic processing (lexical access and contextual integration) of the current word not only successfully but also without a delay. These observations make it hard to just attribute the absence of a prenominal lexical anticipation effect (on the transitive verbs) in the FA dialect to difficulties in lexical access that affect subsequent semantic processing. Instead, the findings in this study suggest that the lack of pre-nominal anticipatory processing in the regional dialects is likely due to the depletion of cognitive resources required for meaning computation that underlies anticipatory lexical processing.

There are two additional reasons why we favor an effortful-meaning-computation mechanism as opposed to a mere automatic-activation account of the lexical anticipation effects. The first is that, in the present study, listeners' long-term knowledge system (e.g., semantic networks) was exactly the same across the three dialect conditions, as all participants listened to the same stimuli spoken in different dialects. Thus, the core semantic information (such as the semantic information decoded directly from the speech signal and that stored in the long-term memory), which is needed to drive automatic activation of semantic associations, are comparable across the home and FA regional dialects. If anticipatory lexical processing in language comprehension was supported only by the automatic activation mechanism, we should have observed the pre-nominal lexical anticipation effect at least in both the home and FA reginal dialect conditions (taking into consideration of possible lexical access difficulties in the UN condition). However, this was not the case, lending further support to the view of non-automatic mechanism(s) for predictive lexical processing.

Secondly, the effortful-meaning-computation account is in line with existing reports that early exposure of home dialect makes the bottom-up perception more efficient and less cognitive demanding (Adank et al., 2009; Cristia et al., 2012 for detailed reviews). In the present study, the bottom-up perception of the home dialect stimuli was indeed easier than that of the two regional dialects, as reflected in the reduced N1 (strong- vs. weak-constraint) effect on the target noun but only in the home dialect condition. We take this as evidence that in listening to the home dialect, the language processing system has more cognitive resources to perform relatively resource-demanding, deeper semantic computations to generate hypothesis of upcoming content, thanks to the native proficiency for easier speech perception. In contrast, when listening to the regional dialects, the bottom-up perception processes may become less efficient, leading to higher processing costs (e.g., Floccia et al., 2006). Consequently, in the case of the FA regional dialect, although lexical meanings could be successfully accessed, the language processing system might not have enough cognitive recourse to expand the currently received semantic information to make predictive inferences. Therefore, no robust anticipatory effect was observed at the pre-nominal verbs in the FA regional dialect condition. The FA condition thus differs from the UN condition, in which participants may not only have increased bottom-up perceptual processing costs but also suffer from higher meaning-level lexical access difficulties in the predictive processing of the target nouns.

It is worth noting that in both Exps 1&2, listeners needed to switch between three kinds of dialects, as utterances from the three dialects were randomly mixed and presented to the listeners. Compared to a design where listeners are confronted with only one dialect, the overall processing difficulty in this study is estimated to be relatively higher. The difficult processing situation of the present study and the processing advantage of early-exposure home dialect might have jointly led to the results that an early effect of anticipatory processing was observed only in the home dialect condition, but not in the familiar or unfamiliar regional dialect conditions. Our results could not inform whether we would observe similar anticipatory processing effects when listeners are exposed to a highly familiar regional dialect in an easier situation, such as just listening to one dialect in one session.

To summarize, the modulating effect of dialectal experience on prenominal anticipatory processing observed in this study suggests that resource-demanding meaning computation is likely to be an important mechanism that underlies anticipatory lexical processing in language comprehension. Needless to say, further evidence is needed from well-designed experiments, which tap specifically into the effects of cognitive resources on predictive processing, for us to understand further the mechanism(s) of lexical prediction.

Conclusions

This EEG study provided new insights into predictive language processing by showing the effects of three dialect listening conditions (i.e. home dialect, familiar regional dialect, and unfamiliar regional dialect) on anticipatory lexical processing. When our listeners encountered stimuli produced in their home dialect or non-native regional dialect, they showed a significant context-induced lexical predictability effect, reflected in the typical N400 reduction over the predictable target noun processing. However, it was only in the home dialect condition that the context-induced ERP effect was observed during the processing of the transitive verb that precedes the target object noun. The correlation between the ERP amplitudes at the verbs and the predictabilities of the following nouns lends strong support to the anticipatory nature of lexical prediction. Moreover, it was also only in the home dialect condition that an N1 effect was observed on the target noun processing, which indicates that the facilitation effect of semantic context on the processing of incoming new words can occur at the early acoustic/phonological stage of speech processing. The N1 and N400 effects observed at the target nouns jointly confirm that language experience modulates the processing stages at which lexical prediction facilitates speech comprehension. The way dialectal experience modulates lexical prediction at the prenominal verbs may be taken as evidence that anticipatory lexical processing involves effortful meaning computation. Our results thus bring furter insights to lexical prediction not only in the temporal sense (i.e., the effect of upcoming predictable nouns on preceding verbs) but also in the spatial sense (i.e., top-down effects from the higher semantic level to the lower acoustic/phonological level).

CRediT authorship contribution statement

Xiaoqing Li: Conceptualization, Methodology, Funding acquisition, Resources, Writing - original draft. **Guiqin Ren:** Funding acquisition, Resources. **Yuanyi Zheng:** Visualization. **Yiya Chen:** Conceptualization, Methodology, Funding acquisition, Writing - revision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Data Availability

The dataset containing the raw and analyzed data is available for public download at the following links:

https://data.mendeley.com/datasets/cf8jy462hn/

draft?a = 8861ba16-d766-4613-a299-8ca1c9b77c3c for raw data of Experiment 1;

https://data.mendeley.com/datasets/j6pnzmfgxz/

draft?a = 35edc86a-149d-4a1c-a95c-ebdf25249f2c for analyzed data of Experiment 1;

https://data.mendeley.com/datasets/r79y9zkpcz/

draft?a = 6067596a-f9d5-44dd-8f62-f83f5533cfbd for raw data of Experiment 2;

https://data.mendeley.com/datasets/g68pmv4wfj/

draft?a = 24c3a5b9-3a5a-4541-9220-f1bc748103a9 for analyzed data of Experiment 2.

References

Adank, P., Evans, B. G., Stuart-Smith, J., & Scott, S. K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 520–529. https://doi.org/10.1037/a0013552.

Adank, P., Van Hout, R., & Velde, H. V. D. (2007). An acoustic description of the vowels of northern and southern standard Dutch II: Regional varieties. *The Journal of the Acoustical Society of America, 121*(2), 1130–1141. https://doi.org/10.1121/1.2409492.

Albareda-Castellot, B., Pons, F., & Sebastián-Gallés, N. (2011). The acquisition of phonetic categories in bilingual infants: New data from an anticipatory eye movement paradigm. *Developmental Science, 14*(2), 395–401. https://doi.org/10.1111/j.1467-7687.2010.00989.x.

Alejandrina, C., Amanda, S., Charlotte, V., Rachel, S., Ann, B., & Caroline, F. (2012). Linguistic processing of accented speech across the lifespan. *Frontiers in Psychology, 3*(479), 1–15. https://doi.org/10.3389/fpsyg.2012.00479.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language, 59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005.

Balconi, M., & Caldiroli, C. L. (2011). Semantic violation effect on object-related action comprehension. N400-like event-related potentials for unusual and incorrect use.

Neuroscience, 197(197), 191–199. https://doi.org/10.1016/j.neuroscience.2011.09.026.

Baggio, G. (2018). Meaning in the Brain. MIT Press. ISBN 9780262038126.

Baggio, G., & Hagoort, P. (2011). The balance between memory and unification in semantics: A dynamic account of the N400. *Language and Cognitive Processes, 26*(9), 1338–1367. https://doi.org/10.1080/01690965.2010.542671.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological), 57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Besson, M., Kutas, M., & Petten, C. V. (1992). An event-related potential (ERP) analysis of semantic congruity and repetition effects in sentences. *Journal of Cognitive Neuroscience, 4*(2), 132–149. https://doi.org/10.1162/jocn.1992.4.2.132.

Brothers, T., Swaab, T. Y., & Traxler, M. J. (2017). Goals and strategies influence lexical prediction during sentence comprehension. *Journal of Memory and Language, 93*, 203–216. https://doi.org/10.1016/j.jml.2016.10.002.

Chambers, C. G., & Cooke, H. (2009). Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(4), 1029. https://doi.org/10.1037/a0015901.

Chirkova, K., & Chen, Y. (2017). Beijing mandarin. In R. Sybesma, W. Behr, Y. Gu, Z. Handel, & J. Huang (Eds.). *Encyclopedia of Chinese language and linguistics* (pp. 275–283). Leiden: Brill.

Clopper, C. G. (2012). Effects of dialect variation on the semantic predictability benefit. *Language and Cognitive Processes, 27*(7–8), 1002–1020. https://doi.org/10.1080/01690965.2011.558779.

Cristia, A., Seidl, A., Vaughn, C., Schmale, R., Bradlow, A., & Floccia, C. (2012). Linguistic processing of accented speech across the lifespan. *Frontiers in Psychology, 3*, 479. https://doi.org/10.3389/fpsyg.2012.00479.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*, 1117–1121. https://doi.org/10.1038/nn1504.

Diaz, M. T., & Swaab, T. Y. (2007). Electrophysiological differentiation of phonological and semantic integration in word and sentence contexts. *Brain Research, 1146*, 85–100. https://doi.org/10.1016/j.brainres.2006.07.034.

Dikker, S., Rabagliati, H., & Pylkkänen, L. (2009). Sensitivity to syntax in visual cortex. *Cognition, 110*(3), 293–321. https://doi.org/10.1016/j.cognition.2008.09.008.

Evans, J. St. B. T., & Over, D. E. (1996). Reasoning and rationality. Hove, UK: Psychology Press. https://doi.org/10.4324/9780203027677.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44*, 491–505. https://doi.org/10.1111/j.1469-8986.2007.00531.x.

Federmeier, K. D., & Kutas, M. (2001). Meaning and modality: Influences of context, semantic memory organization, and perceptual predictability on picture processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(1), 202. https://doi.org/10.1037/0278-7393.27.1.202.

Ferreira, F., & Chantavarin, S. (2018). Integration and prediction in language processing: A synthesis of old and new. *Current Directions in Psychological Science, 27*(6), 443–448. https://doi.org/10.1177/0963721418794491.

Floccia, C., Delle Luche, C., Durrant, S., Butler, J., & Goslin, J. (2012). Parent or community: Where do 20-month-olds exposed to two accents acquire their representation of words? *Cognition, 124*(1), 95–100. https://doi.org/10.1016/j.cognition.2012.03.011.

Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1276. https://doi.org/10.1037/0096-1523.32.5.1276.

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(5), 1461. https://doi.org/10.1037/a0036756.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience, 11*(2), 127. https://doi.org/10.1038/nrn2787.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London, 360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622.

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology, 48*(12), 1711–1725. https://doi.org/10.1111/j.1469-8986.2011.01273.x.

Guillaume, F., Tinard, S., Baier, S., & Dufau, S. (2016). An ERP Investigation of Object-Scene Incongruity. *Journal of Psychophysiology, 32*(1), 1–10. https://doi.org/10.1027/0269-8803/a000181.

Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech compared to reading: The P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia, 38*(11), 1531–1549. https://doi.org/10.1016/S0028-3932(00)00053-1.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304*(5669), 438–441. https://doi.org/10.1126/science.1095455.

Hernández, M., Ventura-Campos, N., Costa, A., Miropadilla, A., & Avila, C. (2019). Brain networks involved in accented speech processing. *Brain and Language, 194*, 12–22. https://doi.org/10.1016/j.bandl.2019.03.003.

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research, 29*(1), 33–56. https://doi.org/10.

1177/0267658312461803.

Hu, M. (1992). Putonghua, the Chinese standard national common speech, and the Beijing vernacular. *Chinese Teaching in the World, 1*, 65–68.

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research, 1626*, 118–135. https://doi.org/10.1016/j.brainres.2015.02.014.

Ito, A., Martin, A. E., & Nieuwland, M. S. (2017). Why the A/AN prediction effect may be hard to replicate: A rebuttal to Delong, Urbach, and Kutas (2017). *Language, Cognition and Neuroscience, 32*(8), 974–983. https://doi.org/10.1080/23273798.2017.1323112.

Ito, A., Corley, M., Pickering, M. J., Martin, A. E., & Nieuwland, M. S. (2016). Predicting form and meaning: Evidence from brain potentials. *Journal of Memory and Language, 157–171.* https://doi.org/10.1016/j.jml.2015.10.007.

Kahneman, D. (2011). Thinking, fast and slow. Farrar. Nov. 2011. c.512p. illus. index. ISBN 9780374275631. https://doi.org/10.1007/s00362-013-0533-y.

Kamide, Y., Altmann, G. T., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and language, 49*(1), 133–156. https://doi.org/10.1016/S0749-596X(03)00023-8.

Kelly, S. D., Ward, S., Creigh, P., & Bartolotti, J. (2007). An intentional stance modulates the integration of gesture and speech during comprehension. *Brain and Language, 101*(3), 222–233. https://doi.org/10.1016/j.bandl.2006.07.008.

Kim, A., & Lai, V. (2012). Rapid interactions between lexical semantic and word form analysis during word recognition in context: Evidence from ERPs. *Journal of Cognitive Neuroscience, 24*(5), 1104–1112. https://doi.org/10.1162/jocn_a_00148.

Kochari, A. R., & Flecken, M. (2019). Lexical prediction in language comprehension: A replication study of grammatical gender effects in Dutch. *Language, Cognition and Neuroscience, 34*(2), 239–253. https://doi.org/10.1080/23273798.2018.1524500.

Klaczynski, P. A., & Daniel, D. B. (2005). Individual differences in conditional reasoning: A dual-process account. *Thinking & Reasoning, 11*(4), 305–325. https://doi.org/10.1080/13546780442000196.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), https://doi.org/10.18637/jss.v082.i13.

Kroll, J. F., & De Groot, A. M. B. (2005). *Handbook of bilingualism: Psycholinguistic approaches.* Oxford University Press.

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience, 31*, 1–28. https://doi.org/10.1162/089892999563724.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4*(12), 463–470. https://doi.org/10.1016/S1364-6613(00)01560-6.

Kutas, M., & Hillyard, S. A. (1980). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology, 11*(2), 99–116. https://doi.org/10.1016/0301-0511(80)90046-0.

Labov, W., & Ash, S. (1997). Understanding Birmingham. In C. Bernstein, T. Nunnally, & R. Sabino (Eds.). *Language variety in the south revisited* (pp. 508–573). University of Alabama Press.

Laszlo, S., & Federmeier, K. D. (2009). A beautiful day in the neighborhood: An event-related potential study of lexical relationships and prediction in context. *Journal of Memory and Language, 61*(3), 326–338. https://doi.org/10.1016/j.jml.2009.06.004.

Lau, E., Almeida, D., Hines, P. C., & Poeppel, D. (2009). A lexical basis for n400 context effects: Evidence from MEG. *Brain & Language, 111*(3), 161–172. https://doi.org/10.1016/j.bandl.2009.08.007.

Lee, C. Y., Liu, Y. N., & Tsai, J. L. (2012). The time course of contextual effects on visual word recognition. *Frontiers in Psychology, 3*, 285. https://doi.org/10.3389/fpsyg.2012.00285.

Lenneberg, E. H. (1967). The biological foundations of language. *Hospital Practice, 2*(12), 59–67. https://doi.org/10.1080/21548331.1967.11707799.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106*(3), 1126–1177. https://doi.org/10.1016/j.cognition.2007.05.006.

Li, X., Zhang, Y., Xia, J., & Swaab, T. Y. (2017). Internal mechanisms underlying anticipatory language processing: Evidence from event-related-potentials and neural oscillations. *Neuropsychologia, 102*, 70–81. https://doi.org/10.1016/j.neuropsychologia.2017.05.017.

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology, 54*(1), 146–157. https://doi.org/10.1111/psyp.12639.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.

Martin, C. D., Thierry, G., Kuipers, J. R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language, 69*(4), 574–588. https://doi.org/10.1016/j.jml.2013.08.001.

McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology Learning Memory & Cognition, 12*(1), 82–91. https://doi.org/10.1037/0278-7393.12.1.82.

Miller, J., Patterson, T., & Ulrich, R. (1998). Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology, 35*, 99–115. https://doi.org/10.1111/1469-8986.3510099.

Mitsugi, S., & MacWhinney, B. (2016). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition, 19*(1), 19–35. https://doi.org/10.1017/S1366728914000881.

Molinaro, N., Conrad, M., Barber, H. A., & Carreiras, M. (2010). On the functional nature of the N400: Contrasting effects related to visual word recognition and contextual semantic integration. *Cognitive Neuroscience, 1*(1), 1–7. https://doi.org/10.1080/17588920903373952.

Molinaro, N., Giannelli, F., Caffarra, S., & Martin, C. (2017). Hierarchical levels of representation in language prediction: The influence of first language acquisition in highly proficient bilinguals. *Cognition, 164*, 61–73. https://doi.org/10.1016/j.cognition.2017.03.012.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology, 24*(4), 375–425. https://doi.org/10.1111/j.1469-8986.1987.tb00311.x.

Nakamura, C., Arai, M., & Mazuka, R. (2012). Immediate use of prosody and context in predicting a syntactic structure. *Cognition, 125*(2), 317–323. https://doi.org/10.1016/j.cognition.2012.07.016.

Nakano, H., Saron, C. D., & Swaab, T. Y. (2010). Speech and span: Working memory capacity impacts the use of animacy but not of world knowledge during spoken sentence comprehension. *Journal of Cognitive Neuroscience, 22*(12), 2886–2898. https://doi.org/10.1162/jocn.2009.21400.

Newman, A. J., Bavelier, D., Corina, D., Jezzard, P., & Neville, H. J. (2002). A critical period for right hemisphere recruitment in American sign language processing. *Nature Neuroscience, 5*(1), 76. https://doi.org/10.1038/nn775.

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., ... Huettig, F. (2018). Large-scale replication study reveals a limit onprobabilistic prediction in language comprehension. *eLife, 7.* https://doi.org/10.7554/elife.33468.

Nieuwland, M. S., Barr, D., Bartolozzi, F., Busch-Moreno, S., Donaldson, D., Ferguson, H. J., ... Ito, A. (2019). Dissociable effects of prediction and integration during language comprehension: Evidence from a large-scale study using brain potentials. *Proceedings of the Royal Society B: Biological Sciences*https://doi.org/10.1101/267815.

Nieuwland, M. S. (2019). Do 'early' brain responses reveal word form prediction during language comprehension? A critical review. *Neuroscience and Biobehavioral Reviews, 96*, 367–400. https://doi.org/10.1016/S0926-6410(02)00193-3.

Obleser, J., Lahiri, A., & Eulitz, C. (2004). Magnetic brain response mirrors extraction of phonological features from spoken vowels. *Journal of Cognitive Neuroscience, 16*(1), 31–39. https://doi.org/10.1162/089892904322755539.

Oostenveld, Robert, Fries, Pascal, Maris, Eric, & Schoffelen, Jan-Mathijs (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011*, 1–9. https://doi.org/10.1155/2011/156869.

Otten, M., & Van Berkum, J. J. A. (2009). Does working memory capacity affect the ability to predict upcoming words in discourse? *Brain Research, 1291*, 92–101. https://doi.org/10.1016/j.brainres.2009.07.042.

Parviainen, T., Helenius, P., & Salmelin, R. (2005). Cortical differentiation of speech and nonspeech sounds at 100 ms: Implications for dyslexia. *Cerebral Cortex, 15*, 1054–1063. https://doi.org/10.1093/cercor/bhh206.

Paul Boersma & David Weenink. (2018): Praat: doing phonetics by computer [Computer program]. Version 5.1.27, retrieved 7 March 2010 from http://www.praat.org/.

Penolazzi, B., Hauk, O., & Pulvermüller, F. (2007). Early semantic context integration and lexical access as revealed by event-related brain potentials. *Biological Psychology, 74*(3), 374–388. https://doi.org/10.1016/j.biopsycho.2006.09.008.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*(4), 329–347. https://doi.org/10.1017/S0140525X12001495.

R Core Team (2018) R: A language and environment for statistical computing, Version 3. 5.0. https://www.r-project.org.

Samuel, A. G., & Kraljic, T. (2009). Perceptual learning for speech. *Attention, Perception, & Psychophysics, 71*(6), 1207–1218. https://doi.org/10.3758/APP.71.6.1207.

Samuel, A. G., & Frost, R. (2015). Lexical support for phonetic perception during non-native spoken word recognition. *Psychonomic Bulletin & Review, 22*(6), 1746–1752. https://doi.org/10.3758/s13423-015-0847-y.

Sanders, L. D., & Neville, H. J. (2003). An ERP study of continuous speech processing: I. Segmentation, semantics, and syntax in native speakers. *Cognitive Brain Research, 15*(3), 228–240. https://doi.org/10.1016/S0926-6410(02)00195-7.

Sohoglu, E., Peelle, J. E., Carlyon, R. P., & Davis, M. H. (2012). Predictive top-down integration of prior knowledge during speech perception. *Journal of Neuroscience, 32*, 8443–8453. https://doi.org/10.1523/JNEUROSCI.5069-11.2012.

Stanovich, K. E., & West, R. F. (2000). Advancing the rationality debate. *Behavioral and Brain Sciences, 23*(05), 701–717. https://doi.org/10.1017/S0140525X00623439.

Sumner, M., & Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language, 60*(4), 487–501. https://doi.org/10.1016/j.jml.2009.01.001.

Szewczyk, J. M., & Schriefers, H. (2013). Prediction in language comprehension beyond specific words: An ERP study on sentence comprehension in Polish. *Journal of Memory and Language, 68*(4), 297–314. https://doi.org/10.1016/j.jml.2012.12.002.

Szewczyk, J. M., & Schriefers, H. (2017). The n400 as an index of lexical preactivation and its implications for prediction in language comprehension. *Language Cognition & Neuroscience, 1–22.* https://doi.org/10.1080/23273798.2017.1401101.

Thornhill, D. E., & Van Petten, C. (2012). Lexical versus conceptual anticipation during sentence processing: Frontal positivity and N400 ERP components. *International*

journal of psychophysiology: Official journal of the International Organization of Psychophysiology, 83(3), 382–392. https://doi.org/10.1016/j.ijpsycho.2011.12.007.

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(3), 443–467. https://doi.org/10.1037/0278-7393.31.3.443.

Van Heugten, M., & Johnson, E. K. (2014). Learning to contend with accents in infancy: Benefits of brief speaker exposure. *Journal of Experimental Psychology: General, 143*(1), 340. https://doi.org/10.1037/a0032192.

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2003). Expecting gender: An event related brain potential study on the role of grammatical gender in comprehending a line drawing within a written sentence in Spanish. *Cortex, 39*(3), 483–508. https://doi.org/10.1111/j.1574-6968.2006.00201.x.

Wicha, N. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience, 16*(7), 1272–1288. https://doi.org/10.1162/0898929041920487.

Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution, 1*(1), 3–14. https://doi.org/10.1111/j.2041-210X.2009.00001.x.