Sound of mind: electrophysiological and behavioural evidence for the role of context, variation and informativity in human speech processing

Nixon, J.S.

Citation

Nixon, J. S. (2014, October 14). *Sound of mind: electrophysiological and behavioural evidence for the role of context, variation and informativity in human speech processing*. Retrieved from https://hdl.handle.net/1887/29299

Author: Nixon, Jessie Sophia
Title: Sound of mind: electrophysiological and behavioural evidence for the role of context, variation and informativity in human speech processing
Issue Date: 2014-10-14

Chapter 4

Context constrains neural activity during speech variant processing: a non-linear model of ERP data

A version of this chapter is in preparation as:
Nixon, J. S., van Rij, Li, X. Q. & Chen, Y. (in preparation). Context constrains neural activity during speech variant processing: a non-linear model of ERP data

Abstract

The phonetic realisation of speech sounds depends on their context, yet most psycholinguistic theories do not account for such phonetic variation. The present study investigated how such phonetic variation is processed by measuring ERP amplitude during a reading aloud task with masked priming. In Beijing Mandarin, Tone 3 usually has a low contour, but preceding another Tone 3 syllable, it has a rising contour. All critical targets had a Tone 2 initial character, which also has a rising contour. In the Contour match condition, primes consisted of two Tone 3 characters (T3 + T3), so the (rising) contour matched the targets, even though the tone category was different. In the mismatch condition, the second character of the prime was another tone (T3 + TX), so prime and target differed in both contour and tone category. ERPs were analysed using Generalised Additive Mixed Modelling, a non-linear model with random effects for subjects and items. Models revealed a complex interaction between prime type and prime and target frequencies. In the mismatch condition, there was relatively little effect of the item frequencies. In the contour match condition, in contrast, there seemed to be a cross-over effect in the prime and target frequencies. When target frequency is relatively high and prime frequency is low, there is reduced negativity. However, when prime and target frequency are both high or both low, there is increased negativity, which suggests competition between prime and target. This suggests that when tonal contour no longer discriminates between prime and target (i.e. in the contour match condition), the a priori probabilities of the prime and target come into play. The conflict that arises between prime and target requires increased processing effort as reflected in greater ERP amplitudes. This difference in the pattern of effects between the Contour match and mismatch primes provides evidence for top-down effects of context on phonological processing of masked primes during reading aloud.

4.1 Introduction

Sub-phonemic processing

Phonetic variation is a fundamental property of speech. Regularities in speech make it possible for speakers to form speech sound categories that distinguish between word meanings, such as between the words ‘pin’ and ‘bin’. In alphabetic languages, sub-lexical processing of speech sounds has been posited to involve activation of strings of phonemes (Dell, 1986, 1988; Foss & Swinney, 1973; Indefrey & Levelt, 2004; W. J. M. Levelt, 2001; W. J. M. Levelt et al., 1999; McClelland & Elman, 1986; Meyer, 1990, 1991; Roelofs, 1999). However, the actual acoustic form of these phoneme categories is far from uniform. It is likely that speakers’ categorisation of sounds reflects the way they are represented in the orthography. For example, in words like ‘spin’, where the first sound is /s/, the second sound is considered to belong to the same sound category /p/ as in the word ‘pin’. But acoustically, the voice onset time falls between the /p/ of ‘pin’ and /b/ of ‘bin’. How this kind of ‘within-category’ variation is processed is not yet well understood. Many current psycholinguistic models, particularly models of speech production and reading aloud, fail to account for processing of context-dependent phonetic variation.

Context effects are well attested during speech perception. Perception of an incoming speech signal is influenced by the semantic context. For instance, processing of meaningless syllables is more similar to real-word processing when semantic information is available from a surrounding sentence (Bonte, Parviainen, Hytönen & Salmelin, 2006). In speech perception, recalibration is a process where the acoustic cues used to discriminate established, native phonemes are shifted after training with contextual or indexical cues. For example, ambiguous acoustic cues are biased towards one or other interpretation by a lexical or visual context or a specific speaker (Dahan, Drucker & Scarborough, 2008; Kraljic & Samuel, 2005, 2007, 2011; Kraljic, Samuel & Brennan, 2008; McQueen et al., 2006; Norris, McQueen & Cutler, 2003). Reinisch, Wozny, Mitterer and Holt (2014) tested whether generalisation of visually guided recalibration of cues contained in meaningless bisyllables depended on phonetic context. Interestingly, they found that recalibration did not occur when the surrounding vowels differed between training and test, suggesting a context-specific utilisation of acoustic cues to distinguish speech contrasts during perception.

Some recent research has begun to address the question of context effects in speech production. Goldrick and Larson (2008) found that speech errors were sensitive to statistical frequencies of syllable positions of features. Errors were less likely to result in fricatives being produced in syllable-final position when 75% of the fricatives in the training were syllable-initial. This shows phoneme processing includes information about the surrounding phonetic context. There is also evidence that overt production and visual processing of speech variants involves multilevel phonological processing. Using the picture-word interference paradigm, Nixon et al. (2014) investigated whether processing of allophonic tonal variants (sandhi words) in Beijing Mandarin involved activation of the tone category (toneme), the context-specific variant (allotone) or both. Experiment 1 revealed that during overt production of sandhi words, both the tone category and the context-specific variant were activated. In Experiment 2, sandhi words were not overtly produced, but were instead visually presented as distractor words superimposed on target pictures. This led to a different time course of effects, compared to overt production. While facilitation from the context-specific variant was stable across both simultaneous and delayed presentation conditions during overt production (Experiment 1), when processed visually as ignored distractor words (Experiment 2) the context-specific variant no longer affected reaction times when presentation of the distractor was delayed. The tone category, in contrast, remained stable for both presentation conditions during visual processing (Experiment 2) but was less robust with delayed presentation during overt production (Experiment 1). This suggests the possibility that there may be differences in processing of sub-phonemic information depending on the task and modality. In particular, during overt production, both the tone category and the context-specific variant are activated early. During visual processing of tonal variants, on the other hand, there seems to initially be activation of the tone category, while it takes time to activate the context-specific variant.

One of the assumptions of priming studies is that activation of overlapping phonological units leads to facilitation of words containing those same units. Few studies have investigated whether priming occurs across speech categories. That is, do similarities in the physical acoustic properties of speech facilitate processing, even if there is no category overlap? Most investigations of sub-lexical phonology have taken the phoneme to be the basic unit of sound. For example, similarity between prime and target is usually measured in terms of phoneme overlap. Recently, a number of studies have provided evidence for sub-phonemic processing of speech sounds (Clayards et al., 2008; Ju & Luce, 2006; McMurray et al., 2009; Mitterer et al., 2011; Newman et al., 2001; Nixon et al., 2014; Nixon, van Rij, Mok, Baayen & Chen, submitted; Nixon, Timmer, Linke, Schiller & Chen, submitted). The contour effect found in Nixon et al. (2014) suggests that acoustic similarity in visual words presented simultaneously with targets is sufficient to facilitate tone processing during speech production, even when there is no category overlap. Whether acoustic similarity in masked primes also facilitates tone processing dur-

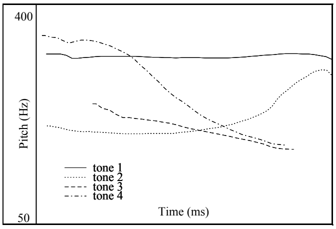


Figure 4.1: Pitch contours of the four tones of Beijing Mandarin

ing reading aloud is not yet known.

Lexical tone processing

Despite the fact that well over half of the world's languages use tone to distinguish between word meanings, a survey of the literature reveals a mere handful of studies on the processing of tone in speech production. As far as we are aware, the present study is the only study to investigate tone processing in reading aloud. Beijing Mandarin has four lexical tones (tonemes). The ptich contours of each of the four tones were shown in Chaper 2, Figure 2.2. They are reproduced here for convenience (Figure 4.1) Characters that have the same segmental syllable can be distinguished by this inherent pitch contour, such as yu2[1] (鱼, 'fish') versus yu3 (雨, 'rain'). In connected speech, Tone 3 (T3) has at least two variants (allotones)[2]. The canonical realisation is the low contour, but preceding another T3 syllable, T3 is realized with a rising contour. This allophonic variant of T3 is known as third tone sandhi (hereinafter, 'T3 sandhi'). Tone sandhi refers to the phenomenon whereby the acoustic realization of a tone is influenced by a neighbouring tone in a particular environment. Importantly for the present study, the contour of T3 sandhi is very similar to another tone, Tone 2. Figure 4.2, reproduced here from Chapter 2, shows the tonal contours of Tone 2, T3 sandhi and the canonical low Tone 3.

The present study has two main objectives. Firstly, it investigates the effects of acoustic contour on reading aloud when prime and target belong to different speech categories. Secondly, it investigates whether the tonal context in which a character occurs affects the relative activation levels of the two alternative tonal variants. One possibility is

---

[1] Mandarin tones are referred to using a number system (Tones 1 to 4). Here, the numeral following the syllable represents the tone number, in this case, Tone 2.

[2] The third tone is sometimes described as 'falling-rising'. However, the rising part of the contour is optional and does not usually occur when there is a following syllable. In addition, the gradient of the fall is very shallow. For these reasons and for simplicity, we refer to the contour as 'low'

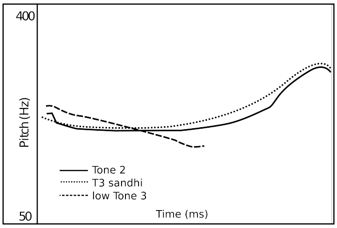


Figure 4.2: Pitch contours of Tone 2, Tone 3 sandhi and canonical low Tone 3

that processing occurs at the speech category level, until the speech preparation stage, when the context-specific articulatory information is activated. A second possibility is that for words containing allophonic variants all variants are activated, regardless of the context in which they occur. Finally, a third possibility is that top-down information available from the surrounding phonetic context contributes to the relative activation levels of the variants, boosting activation of the appropriate context specific-variant. In the former case, since the conditions contain identical initial syllables, we would not expect to see differences between conditions. Only if the actual, context-specific contour is more highly activated in the sandhi (contour) primes than the low-tone (mismatch) primes will we see differences between conditions.

4.2 Method

Participants

Twenty-four native speakers of Beijing Mandarin were paid for participation in the experiment. All participants signed an informed consent form and had normal or corrected-to normal vision.

Materials

Critical targets consisted of 25 two-character Chinese words, of which the initial character was Tone 2 (e.g. 鱼缸, yu2gang1, ‘fish tank’). Each target was preceded by a two-character prime. The initial syllable of each prime was a Tone 3 character, which had the same segmental syllable as the first character of the target. The second character was either Tone 3 (sandhi prime, e.g. 雨水, yu3shui3, ‘rain’) or a different tone (low-tone prime, e.g. 雨衣, yu3yi1 ‘raincoat’). For each target word, the initial character of the prime was identical between prime conditions. When two Tone 3 characters occur together, the first is realised with a rising contour, similar to that of Tone 2. Therefore all prime target pairs differed in terms of the tone category. They either matched or mismatched in the actual contour realisation.

Word frequencies were obtained from Subtlex-CH, a large (46.8 million characters, 33.5 million words) Chinese database based on film subtitles (Cai & Brysbaert, 2010). There was no orthographic overlap (i.e. no shared radicals) between primes and targets. Visual inspection of word frequencies revealed that word frequency was not normally distributed. Therefore, frequencies were transformed using the Johnson transform for normality (jtrans Package version 0.1 in R; Y. C. Wang, 2013) (jtrans Package version 0.1 in R; Wang, 2013).

Table 4.1: Experiment design and sample stimuli

	Contour prime Sandhi (Tone 3 + Tone 3)	**Mismatch prime** Low (Tone 3 + Tone X)
Target yu2gang1 鱼缸	yu3shui3 雨水	yu3yi1 雨衣

Design

The experiment consisted of 200 trials, divided into two blocks of 100 trials, with breaks between the blocks. Each target word was presented twice (once in each prime condition). Two lists were constructed, the order of which was counterbalanced across participants. Lists were pseudo-randomised for each participant. Each block was preceded by three warm-up trials, which were excluded from analysis.

Procedure

Participants were tested individually in a dimly lit, soundproof room, seated approximately 60cm from a 17-inch cathode ray tube computer monitor. A practice session preceded the actual experiment to familiarise participants with the procedure and test the equipment. Stimulus presentation and reaction time data acquisition were conducted using the E-Prime 2.0 software package with a voice key trigger. Participants were instructed to read aloud the words that appeared on the screen as quickly and accurately as possible. All stimuli were presented in black characters on a white background. Each trial began with a fixation cross with jittered presentation time (400-700 ms) to reduce time-induced expectancy waves. A forward-mask of five hash symbols (‘#’) followed for

100 ms, before presentation of the prime for 48 ms. A backward mask (row of hash symbols) was presented for 17 ms to avoid images of the prime remaining on the retina. Finally, the target word was presented for a maximum of 2,000 ms or until the participant response, which triggered the voice key and caused the word to disappear. The experimenter coded incorrect responses and voice key errors in a 1,400 ms interval before the beginning of the next trial. Response time was calculated from the time of target word presentation until the voice key was triggered by the participant response.

4.3 Analysis and Results

Reaction time data: analysis and results

Reaction time data were analysed using linear mixed effects modelling, using the lmer function of the lme4 package (see also Bates et al., 2013; Baayen, 2008; Baayen et al., 2008) in R (R Development Core Team, 2013). Analysis was conducted on the 1,178 data points remaining after stutters, errors, false starts (<1%) and null responses (<1%) were removed. Since error rates were low, no further analyses were conducted on the errors. Inspection of response latency distributions revealed a skewed distribution, which was normalized by logarithmic transformation. Mean response times were numerically longer in the contour match condition (609 ms) than in the mismatch condition (603 ms).

The baseline model was a regression line of log reaction times (log RT), with random intercepts for subjects and target pictures. Prime type was included to investigate the effect of phonetic context on allophonic variants processing and to determine whether congruency in the acoustic contour affects processing of tone during word reading aloud. Johnson transformed prime and target frequencies were also included to investigate whether item frequencies influenced reaction times or interacted with prime type. None of the predictors or their interactions significantly improved model fit. Only target frequency approached significance (p>.06)

Electrophysiological data: recording and pre-processing

The 64-channel electroencephalogram (EEG) was recorded using Neuroscan with electrodes secured in a nylon Electrocap International electrode cap. Electrodes were located at the midline (Fpz, Fz, FCz, Cz, CPz, Pz, POz, Oz) and left and right hemisphere (Fp1, Fp2, AF3, AF4, AF7, AF8, F1, F2, F3, F4, F5, F6, F7, F8, FT7, FT8, FC1, FC2, FC3, FC4, FC5, FC6, C1, C2, C3, C4, C5, C6, T7, T8, CP1, CP2, CP3, CP4, CP5, CP6, TP7, TP8, P1, P2, P3, P4, P5, P6, P7, P8, PO3, PO4, PO5, PO6, PO7, PO8, O1, O2). Eye movements were monitored using additional electrodes placed above and below the left eye and at the external canthi of the left and right eye. These were offline bipolarized to obtain vertical (VEOG) and horizontal electro-oculograms (HEOG). Electrodes placed on the left and right mastoids served as reference points and the GND electrode served as ground. Electrode impedances were kept below 5ω.

The analogue electrophysiological signal was amplified with a band-pass filter between 0.05 and 100 Hz and digitized at a rate of 500 Hz. The digitized EEG was partially processed off-line using Brain Vision Analyzer 2.0. The signal was DC detrended 400 ms before stimulus markers and band-pass filtered from 0.01 to 40 Hz using an inverse discrete wavelets transform. The raw data was then exported and all other pre-processing was conducted in R (R Development Core Team, 2013). The signal was segmented into epochs of 660 ms (160 ms before and 500 ms after stimulus presentation). Each epoch was baseline corrected on the 160 ms before target onset using the baseline function of the eRp package (version 0.9.8.11; Tremblay, 2013a). Blinks and other ocular artefacts were corrected based on vertical and horizontal EOGs using independent component analysis (ICA) with the icaOcularCorrection package in R (Tremblay, 2013b). Figure 4.3 shows the grand average wave form for the mismatch and contour prime conditions at nine electrodes. In all figures, positive is plotted up.

Electrophysiological data: analysis and results

EEG data were analysed using Generalized Additive Mixed Modeling (GAMM; Wood, 2006) see also Tremblay & Baayen, 2010) using the mgcv package 1.7-28 (Wood, 2011) conducted in R (version 3.0.1; R core team, 2013; www.r-project.org). GAMM is a type of Generalised Linear Model (GLM) that uses non-linear smooth functions to model linear predictors. The method used for this is penalized iteratively reweighted least squares (PIRLS; see Wood, 2006, for details). PIRLS determines the optimal linear and/or non-linear equation for avoiding both over-fitting and over-generalizing of the model.

Due to the high computational cost of modelling each electrode, analysis was conducted on a reduced set of 24 electrodes with an even distribution over the scalp (Figure 4.4). Decompositional models were initially run to determine which individual predictors and interactions influenced EEG amplitude. Auto-correlation (i.e. rho) parameters were determined for each model. The full model included trends over time as random effects for participants and items. A smooth of time was included in the model to examine the changes in amplitude over the course of each trial. Our main predictor of interest, prime type, was included in order to

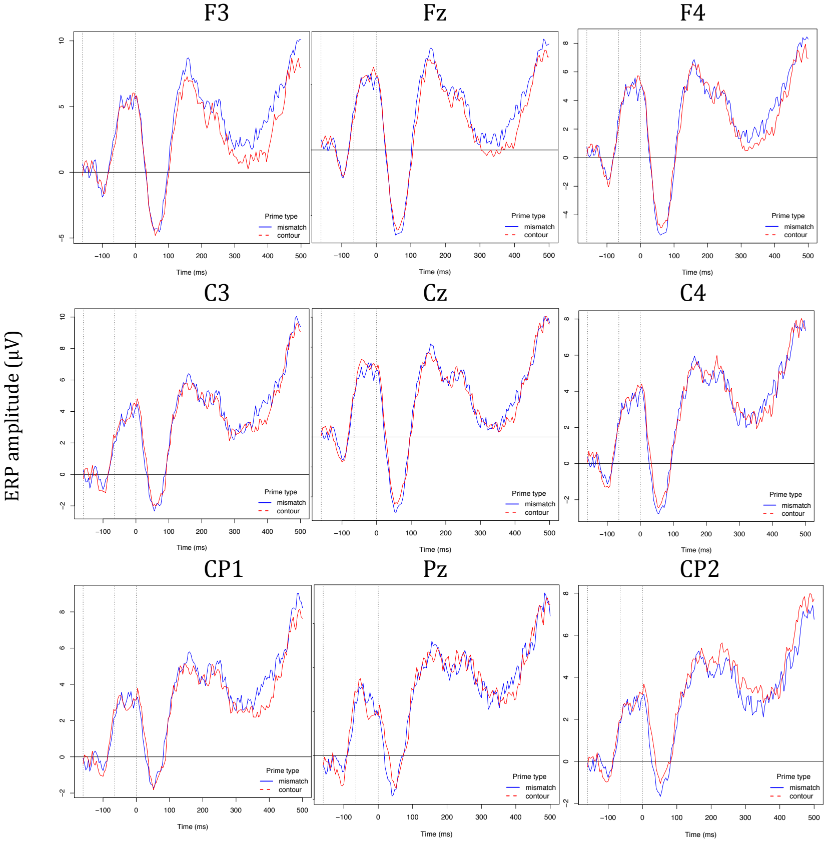


Figure 4.3: Average ERP signal for Contour and Mismatch primes at nine electrodes

determine whether the phonetic context of the following tone affects the processing of Mandarin tone variants and whether cross-category acoustic similarity affects processing during reading aloud. Prime condition was coded as a binary variable, so that in each sample, prime was either low-T3 (0) or sandhi (1). Johnson-transformed prime and target frequencies were also included in the model to investigate whether individual word frequencies affected amplitude of the signal or interacted with the effect of prime type. All two- and three-way interactions were included for each prime type. Model summaries showed significant higher-order interactions between prime type, prime frequency and target frequency over time.

These interactions were then investigated by running separate tensor

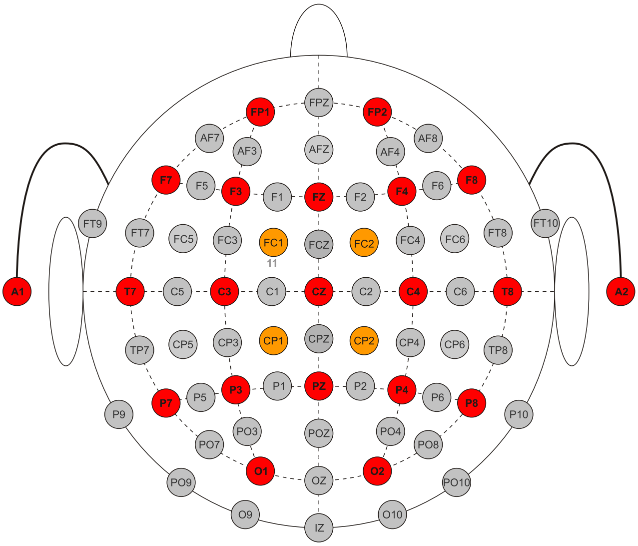


Figure 4.4: Electrode map: models were run for channels marked in red and orange

models for each individual electrode. Rho parameters were determined for and included in each model for each channel. Model comparisons were run separately for each channel using the compareML function (version 2.0; Van Rij, 2014) to determine which predictors contributed to model fit. The compareML function performs a chi-square test on the fREML scores of each model, taking into account the degrees of freedom. Results of the model comparisons are shown in Appendix A. Appendix B shows the model summaries for the best-fit model for each individual electrode.

Random effects

The best-fit model includes trends over time as random effects for participants and target items. The random wiggly curves are shown in Figure 4.5 for participants (left panel) and items (right panel) for the Fz elec-

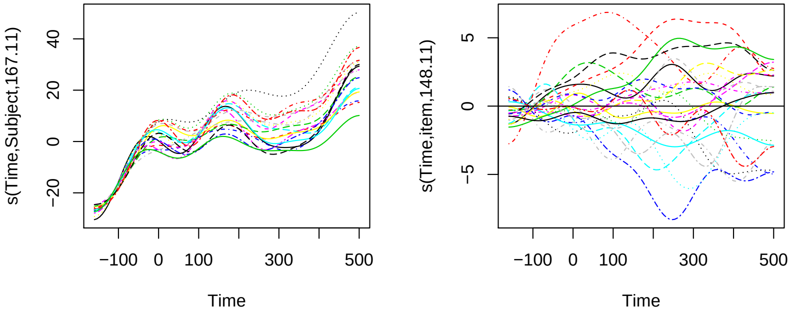


Figure 4.5: Random wiggly curves over time for subjects (left panel) and items (right panel) for the Fz electrode

trode. The figures show the by-subject and by-item adjustments to the amplitude over time.

Fixed effects

Model comparisons (Appendix A) show that model fit was significantly improved for almost all channels by inclusion of a predictor of prime frequency over time (model 2 versus model 1). The models were further improved by an interaction between prime type, prime frequency and target frequency over time (model 3 versus model 2). This 3-way interaction effect over time was consistent across almost the entire scalp (Appendix B).

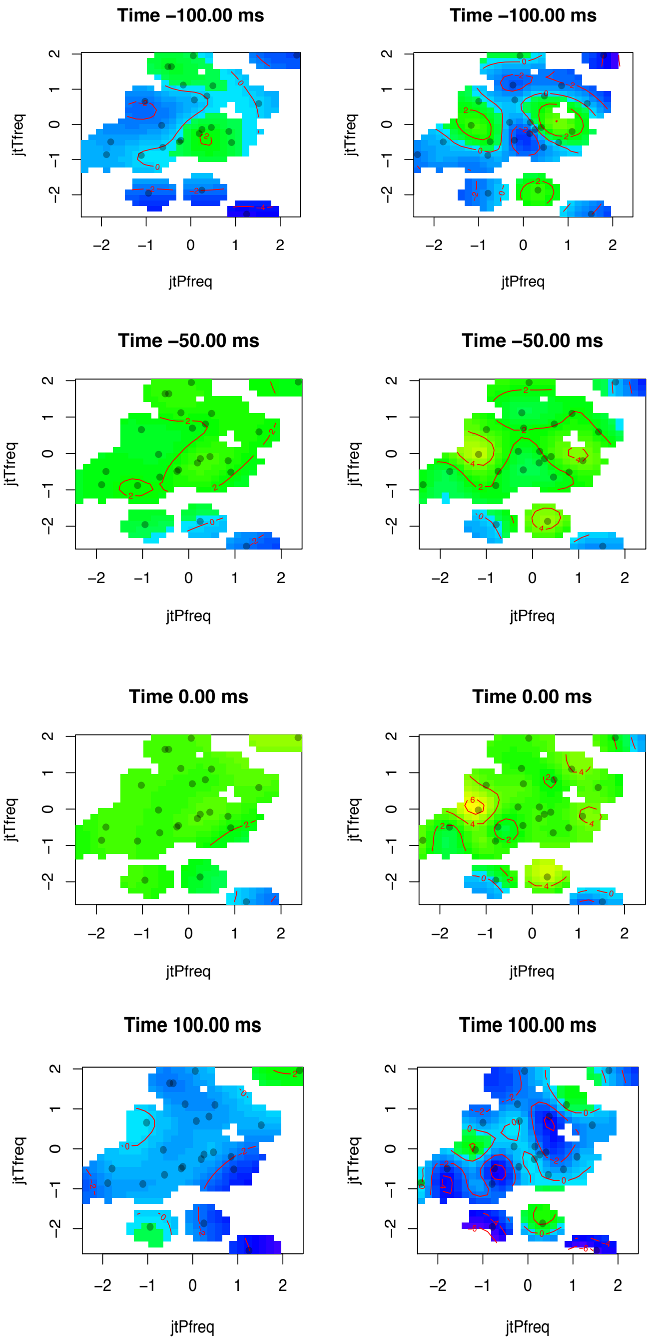
Figure 4.6 shows the interaction between prime frequency and target frequency over a series of time points (-100, -50ms, 0 ms, 100 ms, 200 ms, 300 ms, 400 ms) for trials with the control prime (left panels) and contour match prime (right panels) for electrode Fz. Yellow indicates relatively more positive amplitude; blue indicates more negative amplitude, green is at the intercept. On the x-axis is prime frequency and on the y-axis is target frequency. The panel rows show the different time points from -100 ms in the top row to 400 ms in the bottom row. The grey dots indicate the individual items.

Consistent with the model comparisons and model summaries, the plots show that the two prime types elicit a different pattern of effects. In the mismatch prime condition (left panels), amplitude is relatively flat across prime and target frequencies over time. In the match condition, in contrast, there is a much greater effect of prime and target frequency over time. At -50 ms and 0 ms, the distribution is quite flat in both conditions; however, at 100 ms, there is greater negativity in the match condition, indicated by the dark blue pools. This is near the peak of the first negative-going waveform. By 200 ms, following the peak of first positive-going waveform, the plots show that the amplitude for both prime types is becoming more positive. In the mismatch prime condition, the amplitude is still relatively flat across frequencies, while in the match condition, positive and negative peaks begin to emerge. These positive and negative peaks continue to increase at 300 ms and 400 ms. Generally speaking, there seems to be more negativity with higher-frequency primes and lower-frequency targets. However, this does not account for all the variance. This point will be returned to in the Discussion.

4.4 Discussion

In a reading aloud task with masked primes, electrophysiological measurements (EEG) and reaction times were used to investigate how phonetic context affects processing of allophonic speech variants. Participants read aloud two-character Mandarin words preceded by briefly presented (48 ms) two-character masked primes whose initial character always differed from the target in terms of tone category, but either matched or mismatched the tone contour (that is, the actual acoustic realisation). The lexical phonetic context Ðspecifically, the tone of the following character - determines whether the tone of the initial character is realized in its canonical (low-tone; mismatch condition) form or in its *sandhi* (rising; contour match condition) form.

No significant differences were found in reaction times between conditions. However, the EEG analysis revealed significant between-condition effects. In the mismatch condition, amplitude of the EEG signal was relatively flat across prime and target word frequencies over time. However, in the contour match condition, there was more negativity on the first negative-going waveform at 100 ms and on the second negative-going waveform, particularly from 300 ms to 400 ms. This effect was modulated by prime and target frequency. There was generally more negativity for lower-frequency targets and higher-frequency primes, but this pattern was not consistent for all items. There were also peaks of positivity at central values. Note that in traditional EEG analysis using ANOVA, only the main effect of prime over time would be analysed, since data are averaged over conditions and analysis of by-item characteristics is not possible. In fact, with the present data, if prime and target frequency information is excluded from the models, prime type is still significant. However, since prime and target frequency contributed to model fit, they were retained in the models. Therefore, it is clear


Time −100.00 ms
Time −100.00 ms
Time −50.00 ms
Time −50.00 ms
Time 0.00 ms
Time 0.00 ms
Time 100.00 ms
Time 100.00 ms
jtTfreq
jtPfreq
−2
−1
0
1
2

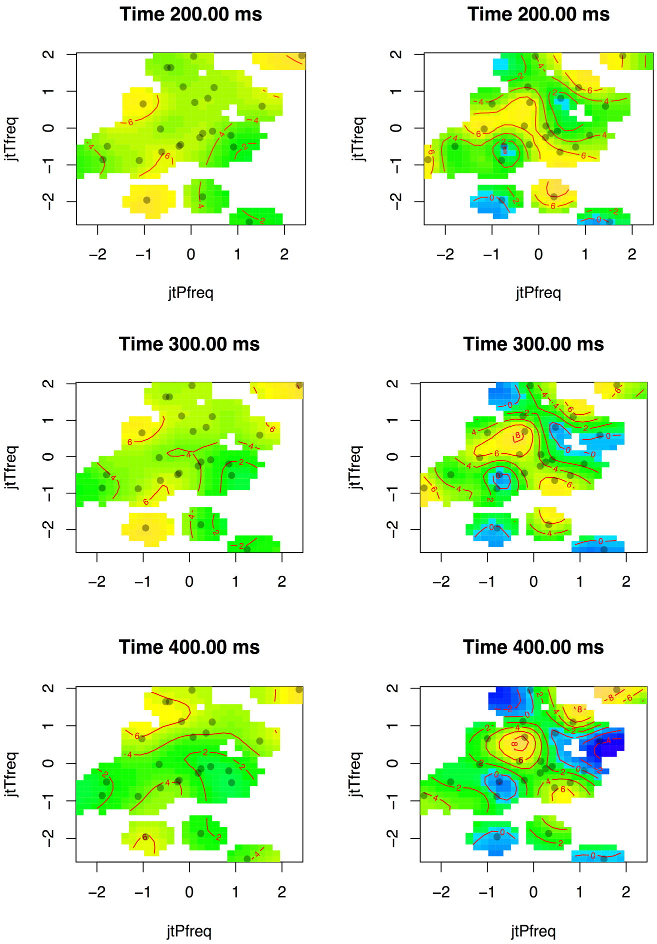


Figure 4.6: Interaction between prime frequency and target frequency at time points -100 ms, -50ms, 0 ms, 100 ms, 200 ms, 300 ms, 400 ms for trials with the control prime (left panels) and contour match prime (right panels) at electrode Fz. Yellow indicates relatively more positive amplitude; blue indicates more negative amplitude; green is at the intercept. Prime frequency is on the x-axis; target frequency is on the y-axis. The panel rows show the different time points from -100 ms in the top row to 400 ms in the bottom row. The grey dots indicate the individual items.

that processing is significantly modulated by prime type, indicating that there is greater activation of the context-specific contour in the contour match condition, compared to the mismatch condition. This supports the hypothesis that processing of phonetic variants is context-specific, i.e. that the surrounding phonetic context constrains activation of the alternative variants. Effects were seen on both the first and second negative-going waveforms. The early time window is associated with sub-phonemic processing of masked primes in silent reading and reading aloud (Ashby et al., 2009). The later time window is associated with sub-lexical phonological processing (e.g. Carreiras, Perea, Vergara & Pollatsek, 2009; Grainger, Kiyonaga & Holcomb, 2006).

What is not yet entirely clear is the role that prime and target frequency play here. It may be that the current analysis does not include the correct measures. For example, it may be that the results are better predicted by bigram frequencies or other co-occurrence measures, rather than simple word frequency measures. There is one other important point to be made here. Although amplitude is comparable between conditions at -50 ms and 0 ms, the plot for -100 ms shows that the model is showing differences between conditions at this early time point. In fact, this should not be possible, because neither prime nor target have been seen yet at this time. In the present study, there were 25 items per prime condition. It may be that this number of items was too small relative to the effect size and that effects may be driven by individual item characteristics. Further work is need to disentangle the precise role that item frequency plays.

Nonetheless, the importance of these frequency measures in processing the target depends on whether or not there is contour overlap between prime and target. This result is reminiscent of previous findings in the tone processing literature. In a Mandarin implicit priming study, J. Y. Chen, Chen and Dell (2002) showed that there was significant priming from overlap of the segmental syllable + prime, and less but still significant priming from the segmental syllable only (when tones differed), but there was no facilitation from tone-only overlap. Similar effects have been found in Cantonese word production (Wong & Chen, 2008). Matching tones did not facilitate processing if the segmental syllable differed. Similarly, in the present results, frequency measures seem to only come into play when there is contour overlap.

The present results extend findings from a previous study, which found evidence for multi-level processing of speech variants (Nixon et al., 2014). This picture-word interference study showed that processing occurs both at the tone category-level and at the level of the context-specific tone variant during overt speech production. Naming of target pictures with sandhi names was facilitated both when target and distractor matched in tone category, but mismatched in tone contour (toneme condition), and when target and distractor matched in tone contour, but mismatched in tone category (contour condition), compared to controls, which mismatched in both tone category and contour. A second experiment suggested that multi-level processing also occurs when tone variants (sandhi words) are not overtly produced targets, but instead visually presented, ignored distractor words superimposed on targets. The present study shows that automatic context-specific processing can also occur during reading aloud.

The direction of effects was counter to predictions. Based on the earlier study (Chapter 2 of this thesis, Nixon et al., 2014), we expected a facilitatory effect when prime and target matched in contour. Here we see mostly a larger ERP amplitude in the match condition. There are a number of factors which may explain the different direction of effects between the two studies: inhibitory effects in the present experiment, compared to facilitatory effects in the previous study. Firstly, the present study used a different manipulation of the tonal conditions compared to the picture-word interference (PWI) study. The PWI study compared contour match conditions with an unrelated tone (Tone 1 or Tone 4). Although the present results show that context constrains the degree to which the allophonic variants are activated, it is likely that the context-inappropriate allophonic variant receives at least partial activation. This may lead to a reduction in the relative facilitation between conditions. Secondly, the presentation time of the prime relative to the target differed between studies. In the PWI study, the contour-congruency effect was sensitive to the manipulation of stimulus onset asynchrony (SOA). The contour was beneficial only when presented simultaneously with the target. When presentation was delayed (83 milliseconds after the target), there was no facilitation effect from the contour, although there was still facilitation from the tone category. This suggests that, with visually presented stimuli, the sandhi contour may take longer to generate than the tone category. In the present experiment, the prime was presented 65 ms prior to the target word, so lexical retrieval of the prime should have progressed further at the time of target onset, compared to the PWI study. It may be that at early stages, similarity in the acoustic realisation is beneficial to lexical retrieval. However, as lexical access progresses, activation of the matching contour leads to suppression (inhibition) of the competing Tone 2 category.

In summary, the present results provide evidence for automatic retrieval of sub-phonemic information in visually processed masked prime words. Although prime and target belonged to different tone categories, a match in the actual physical contour led to differences in the amplitude of the encephalogram, compared to when there was no contour match between prime and target. Interestingly, counter to predictions, contour congruency led to greater negativity over time, depending on prime and target frequency. This may reflect an interference effect due to competition between the mismatching tone categories, following the matching contour. Finally, the finding that target word frequency influences neural activity and interacts with other predictors suggests that it is useful to include individual item characteristics in ERP studies of language processing.

Appendix A Model Comparisons

Predictor Channel	Fp1	Fp2	Fpz	F7	F3	Fz	F4	F8	FC1	FC2	CP1	CP2	T7	C3	Cz	C4	T8	P7	P3	Pz	P4	P8	O1	O2
Interaction of target and prime versus target only	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
Interaction of target and prime versus prime only	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
Predictor of target frequency	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Predictor of prime frequency	***	***	***	***	*	***	***	***	***	***	ns	***	***	***	***	***	ns	ns	***	*	***	ns	**	***

Appendix B Model Summaries

Channel	Fp1	Fp2	Fpz	F7	F3	Fz	F4	F8	FC1	FC2	CP2	T7	C3	Cz	C4	P3	Pz	P4	O1	O2
Predictor																				
Time	***	***	***	***	***	***	***	***	***	***	***	*	***	***	***	***	***	***		***
jtPfreq				.	***	**	*			**	**	.	**		*			*	*	***
jtTfreq									*	*				.	*			***		.
Time:prime	**				*		*	***				**							***	
jtPfreq):prime		***	***	***		***	**	***	***	***	***	***	***	***	***	***	***	***	***	***
jtTfreq):prime		***	***	***			***	***	***	**	***	***	**	***	*	***		***		***
Time,jtPfreq		**			*						*		.						*	***
Time,jtTfreq	**	***		*		*				.	***		**		*	***	***	***		
jtPfreq,jtTfreq					*	.		*				.	.			**		.		
Time,jtPfreq):prime	***	*	**	.	***	***	***		***	***	***		**	**	***				*	
Time,jtTfreq):prime	.	***	***	***		**	*			*	*		*		*	**	.	*	.	
jtPfreq,jtTfreq):prime	**	.	*	***	***	***	***	**	**	***	***	**	***	***	***	**	**	***		
Time,jtPfreq,jtTfreq	*	***		.	**	*	.	***	***	***	*	**		**			*	*	***	***
Time,jtPfreq,jtTfreq):prime	***	***	***	***	***	***	***	***	***	***	***	***	*	***	***	**	.	***	***	**