THE ONLINE PROCESSING OF ANTICIPATORY TONAL INFORMATION IN STANDARD CHINESE BY NATIVE AND NON-NATIVE LISTENERS

Hana Nurul Hasanah [1,3], Qing Yang [1,2], Yiya Chen [1,2]
[1] Leiden University Centre for Linguistics, [2] Leiden Institute for Brain and Cognition, [3] Chinese Studies Universitas Indonesia
h.n.hasanah@hum.leidenuniv.nl, q.yang@hum.leidenuniv.nl, yiya.chen@hum.leidenuniv.nl

ABSTRACT

This study investigated whether and how anticipatory tonal coarticulatory cues are utilized during online speech processing. Native speakers of Standard Chinese (SC) and Indonesian learners of SC participated in an eye-tracking experiment with the Visual World Paradigm (VWP). They were asked to identify two-syllable nonce words that were printed on a screen while listening to auditory stimuli that had varying levels of appropriate anticipatory tonal cues. The eye movement data revealed that (a) both groups showed some sensitivity to anticipatory tonal information when selecting a target item; (b) native listeners of SC demonstrated a predictive effect of tonal processing when the anticipatory tonal cues in the first syllable were strong, but this effect was not present for Indonesian-speaking SC learners.

Keywords: online speech processing, anticipatory tonal information, eye-tracking

1. INTRODUCTION

Listeners are known to use cues of anticipatory coarticulation in an incoming speech stream (e.g., [1, 2]) to speed up spoken word recognition (e.g., [3, 4]). Evidence, however, has been drawn mainly from segmental coarticulation in West-Germanic languages like English and Dutch. There is much less research on the possible utilization of anticipatory tonal cues in languages with lexical tone, such as Sinitic varieties. We know that in Standard Chinese (SC), tones show anticipatory coproduction in connected speech (e.g., [5] for bisyllabic context and [6] for trisyllabic context). Studies of real-time spoken word recognition in SC have revealed the important role of tonal (in addition to segmental) information in listeners' recognition of a target word before the auditory stimulus is entirely unmasked [7–10], and even for ambiguous acoustic signals with gradient tonal manipulation (e.g., for native SC listeners: [11]; for both native and non-native (English) listeners:[12]).

To date, however, only [13] have investigated the use of anticipatory tonal cues in SC spoken word recognition. Using the VWP, they examined listeners' looks to target-competitor words in Tone3 (T3) which differed in segments. They found that native SC listeners could utilize the T3 sandhi information to anticipate an upcoming T3 word. Note that tonal variation, including the anticipation cues, can be broadly classified as due to phonological alternation, known as tone sandhi, and to phonetic alternation, known as coarticulation (even though their boundary can be difficult to draw) (see [14] and references therein). Tone sandhi typically induces salient pitch changes, e.g., the T3 sandhi in SC [15]; its magnitude of tonal anticipatory effect is larger than coarticulation-induced anticipatory effects. It remains unclear whether subtle anticipatory cues due to tonal coarticulation are sufficient for listeners to speed up lexical tone access and to what extent coarticulation effects differ from tone sandhi effects. It is also unclear how general listeners' ability is to use anticipatory tonal information for speech processing and to what extent L1 speaker and L2 learners of a tonal language may differ in their utilization of anticipatory pitch cues.

To address these questions, the current study was designed to understand the role of different types of anticipatory tonal information in the online speech processing of both native and non-native listeners of SC. We chose Indonesian learners of SC which helps to generalize our findings to non-tonal L2 learners of SC whose first language is not a commonly studied Germanic language. We implemented the Visual World Eye-tracking Paradigm with two printed nonce words (NWs) (see [16] for a review of the paradigm). Listeners saw the nonce words on a screen while listening to auditory stimuli. We examined their eye fixations towards the printed stimuli.

The printed stimuli differed in the tone of the second syllable; the auditory stimuli either had coarticulatorily appropriate or inappropriate (i.e., matched vs. mismatched) anticipatory pitch cues. Comparing listeners' responses between the two types of cues is not only a classic approach to examine the benefit of anticipatory cues in perception (e.g., [1]) but also recommended for the investigation of anticipatory tonal cues in spoken word recognition [17]. We included two tonal contexts that differed in the magnitude of tonal anticipatory cues: strong magnitude due to tone sandhi (Tone3; T3), to replicate the findings in [13], and weak magnitude due to coarticulation (Tone1; T1). Our goal was to compare whether degrees of anticipatory coarticulation affect how the coarticulatory cues are used.

Assuming listeners of SC use anticipatory tonal cues in online speech processing, we expected that when an auditory stimulus contains a mismatched tonal cue, there would be a delay in listeners' fixation on the target and consequently, temporarily decreased looks to the target. Furthermore, a strong anticipatory cue (T3 context) would amplify the cue effect and speed up listeners' recognition of the spoken nonce word, relative to the weaker coarticulatory cue (T1 context). Last, comparisons in the gaze patterns of native vs. non-native listeners would shed light on the general ability of listeners to utilize anticipatory tonal information for online speech processing and the role of L1 in modulating the use of such cues.

2. METHOD

2.1. Participants

Thirty-two native SC speakers (mean age: 26.3, SD: 3.6, 26 females) and thirty-one Indonesian L2 learners of SC (mean age: 20.9, SD: 0.7, 27 females) participated in this study.[1] Two SC speakers also reported speaking another Northern Mandarin variety. All L2 learners had learned SC for 2.5 to 3.5 years (2.5 years = 14 learners) and achieved ±2500 words (Intermediate level).[2] All participants reported normal to corrected vision and no speech and hearing disorders history.

2.2. Design

Table 1 illustrates the 4 x 2 within-subject design.

Factor	Level
Trial condition	1. critical 2. baseline
Appropriateness of coarticulatory cue	1. matched cue 2. mismatched cue
Magnitude of the anticipation (Tone type)	1. strong (T3) 2. weak (T1)
L1 background	1. tonal language 2. non-tonal language

Table 1: Experimental factors

Apart from the baseline and critical trials (66 each), fillers were added (164 trials) to balance the occurrence of items in all levels of trial condition (Appendix 1). These 296 items were presented in four blocks. We also included one practice block of 10 trials. Trial randomization was implemented within each block, and the position of target-competitor items was randomized and counterbalanced. Each item only appeared once in each block in either matched or mismatched cues.

2.3. Visual stimuli

The target NWs carried T3 or T1 followed by T3 (T3T3; T1T3) and had either a sonorant or a stop (/b/ or /d/) onset in the 1st syllable. As visual stimuli, they were typed in black KaiTi font and appeared with either a phonological competitor or a distractor. All items were controlled for visual complexity, measured in their number of strokes [$F(2,96)=0.731$, $p=0.484$], log morphemic frequency according to the SUBTLEX-CH corpus [18] [$F(2,96)=1.951$, $p=0.148$], and phonological neighborhood density based on the Chinese Lexical Database [19] [$F(2,96)=0.856$, $p=0.428$]. Furthermore, the goodness of the NWs was judged by fourteen native speakers of SC on a 4-point scale (1 meaning "the word is not a real word in SC" and 4 meaning "the word is a real word in SC"). We only included pairs where each member had an average value of less than 2.

2.4. Auditory stimuli

Bisyllabic NWs were recorded at the Phonetics Lab of the Leiden University Centre for Linguistics by one male native SC speaker, who was born and raised in Beijing. The NWs were produced in a carrier sentence, "*wo3 nian4* ___", and recorded in a randomized order three times. Two tokens of each word were selected to create spliced stimuli. The selection was based on accuracy and clarity in the patterns of anticipatory coarticulation and sandhi alternation (see Appendix 2). In total, 33 NW triplets were included in this experiment.

To manipulate the tonal coarticulatory information, the T3 in the second syllable was spliced onto the first syllable of another word with either an identical tonal context (i.e., matched cue) or a different tonal context (i.e., mismatched cue). Furthermore, the duration of the first syllable was normalized to 330ms, which was the mean duration of the first syllable of target-competitor NW pairs. The manipulation and normalization were done in Praat using customized scripts [20], [21]. The filler items were also manipulated in this way.

2.5. Procedures

Our participants were tested in a dim room and seated on a height-adjusted chair. A chin-rest was placed in front of an Eyelink Portable DUO (version 6.10.01) eye-tracker, which was situated central-horizontally from a 23-inch monitor. The task was created and presented using SR Research Experiment Builder (version 2.2.245) with auditory stimuli played using a soundcard (M-Track 2x2M) over a Beyer DT-770 dynamic headphone. The visual display was left-right symmetric on-screen against a white background.

Prior to the eye-tracking task, there was a training session to familiarize participants with the intended coarticulatory tonal variations and ensure they knew the pronunciation of all NWs stimuli. They listened to a list of real bisyllabic words, then asked to read aloud the NW stimuli, which was recorded. Only participants who produced all items as intended could begin the eye-tracking task. Native and non-native participants differed significantly in this session, with L1 speakers spending ±25 minutes and L2 learners 50-120 minutes to complete the task.

We included a 9-point calibration and validation procedure before the eye-tracking recording. Each trial began with a central fixation cross (500ms) with an invisible boundary for a drift check. Participants' gaze should be within the boundary to allow visual stimuli to be presented for 1000ms. Otherwise, recalibration and validation were enforced. After an auditory stimulus was played, participants were instructed to select the corresponding printed NW using a mouse click. For the entire session, written instructions were given in the native language of the respective groups, and participants were allowed to rest during the block transition. Nobody took a break of more than 5 minutes.

2.6. Data analysis

Data containing fixation samples and proportions were extracted in 20ms bin length using the Time Binning Report in SR research Data Viewer (version 4.1.1). We only included trials identified correctly by participants and fixation samples that fell on the target or competitor area region. The starting time-point of analysis was 200ms post stimulus onset to account for the oculomotor delay for programming a saccade [22]. While the ending time-point was where, approximately, the proportion of looks to target had reached the maximum, i.e., 1000ms (native listeners) and 1200ms (non-native listeners). For this report, statistical modeling for native and non-native data was performed separately to reduce model complexity. Patterns of fixation proportions across conditions over time for each group were modeled using generalized additive mixed modeling (GAMM) (*mgcv* package version 1.8-41) [23], and model plots were done in the *itsadug* package (version 2.4.1) [24] in R (version 4.2.0) [25]. The ineffectiveness of binomial distribution to account for autocorrelation led us to the continuous response distribution and we logit transformed the fixations proportions ([26, 27]).

Different smoothing parameters were implemented for the effects of our interest, which allowed significance tests from the conditional effects to be reported by pairs of parametric and smooth terms. The final model for each listener group included time as a covariate, trial condition (Trial), appropriateness of coarticulatory cue (Cue), the magnitude of anticipation (Tone) and their interactions as fixed effects, and by-participant factor smooth as a random effect. We also incorporated an AR(1) error model to reduce autocorrelation in the residuals. All models were fitted using a scaled-t family to alleviate the non-normal distribution of the residuals. *P*-values were adjusted using the Holm-Bonferroni method to control the family-wise error rate [28]. Model summary outputs of fixed effects are presented in Appendix 3.

3. RESULTS

Proportions and trajectories of target fixations varied over time, with the native listeners demonstrating a higher average proportion of fixations (POF) than the non-native listeners. The parametric and smooth terms in each group's model summary showed a significant effect of Trial ($p<0.01$). Specifically, the estimated average POF in critical trials was lower than in baseline trials and the difference smooth term contrasting the shape of fixation between the two conditions was far above 0. Last but not least, we also found a significant interaction of Trial and Tone ($p<0.05$), but only for native listeners.

As we zoomed into the critical trials, we found a common pattern shared by native and non-native listeners, i.e., higher target POF means in the matched-cue condition than in the mismatched-cue condition (visually depicted as a concave shape in **Fig. 1**). However, it was only in the T3 context that the difference was significant. Moreover, for the native listeners, the difference smooth term also reported a significant effect of Cue ($p<0.05$). Fixation trajectories between matched- and mismatched-T3 started to differ significantly at around 200ms before the offset of the 1st syllable (see **Fig. 1*a***). These patterns were not observed in the T1 context (see **Fig. 1*b***), which confirmed the significant effect of Tone ($p<0.05$) reported in the model summary. This difference between T3 and T1 context suggests that degrees of anticipatory coarticulation affected how the coarticulatory cues were used.

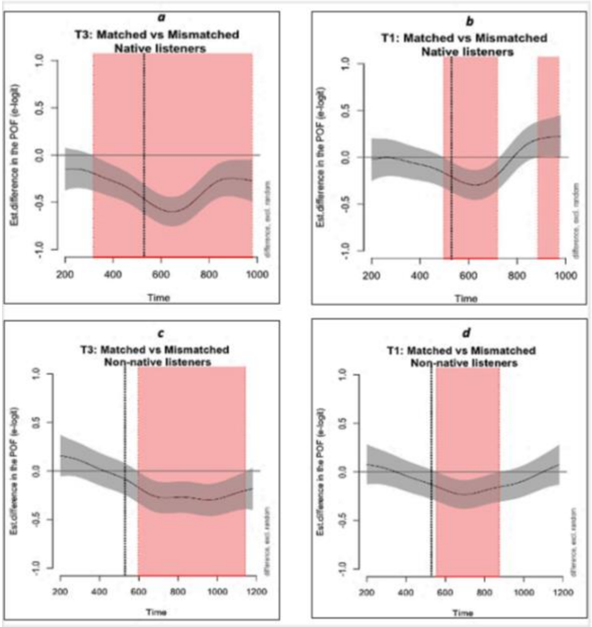


Fig. 1: Estimated difference smooth between matched- and mismatched-cue conditions in critical trials. Top panels: native listeners' T3 (***a***) and T1 (***b***). Bottom panels: non-native listeners' T3 (***c***) and T1 (***d***). The vertical shade indicates that a horizontal confidence band significantly differs from zero, that is, the interval of an effect. The vertical dotted line represents the first syllable offset (530ms).

As for the non-native listeners, the main effect of Cue was absent; looks between the matched vs. mismatched cue conditions became significantly different only after the offset of the 1st syllable, even when the anticipatory cues were salient (see the start of the vertical shade in **Fig. 1***c* and **1***d*). The absence of the effect of Cue, and also Tone, indicates the difference between native and non-native listeners in their utilization of anticipatory tonal cues in online speech processing.

4. CONCLUSION AND DISCUSSION

Our results found significant differences between the baseline and critical trials, suggesting that the presence of phonological competitors in the critical trials significantly inhibited both native and non-native listeners from fixating on the target. Moreover, for the critical trials, both groups were slower in fixating on the targets and showed fewer looks when the auditory stimuli contained inappropriate anticipatory tonal cues. This pattern supports the prediction that listeners are sensitive to mismatched tonal coarticulatory cues, as evidenced by a delayed onset of target fixation and temporarily decreased looks to the target in a cue-mismatched condition.

Furthermore, the magnitude of anticipatory cues modulated listeners' utilization of anticipatory tonal information in online speech processing. In the T1 coarticulation context, neither group of listeners was able to use the cue to anticipate the upcoming tone in a timely manner (**Fig. 1*b*** and **1*d***). In the T3 sandhi context, native listeners were able to utilize the anticipatory cue at approximately 200ms before the end of the first syllable (**Fig. 1*a***), suggesting a predictive effect of the tonal cue. Our findings thus confirm the perceptual advantage of anticipatory tone sandhi cues for native listeners reported in [13].

Regarding non-native listeners, though they demonstrated a longer interval of difference effect in the T3 condition than the T1 condition (**Fig. 1***c* and **1*d***), they could not use the T3 sandhi anticipatory cue predictively, like the native listeners (**Fig. 1*a*** and **1***c*). Another difference between native and non-native listeners is that non-native listeners had fewer target fixations for all conditions. This may be due to their less efficient (therefore slower) decoding for lexical or sentence processing [29, 30], and in this study, less efficiency in utilizing the tonal cues. The longer training session for the non-native group also suggests it required more cognitive effort to process the novel (nonce) words.

To sum up, both groups utilized anticipatory tonal cues during online speech processing as they fixated on a target more quickly in the cue-matching condition than in the cue-mismatching condition. However, non-native listeners did not show an immediate use of T3 sandhi cue to anticipate an upcoming tone, as native listeners did. The observed differences between native and non-native listeners in their ability to utilize anticipatory tonal cues confirm that the tonality of L1 prosody significantly impacts their processing efficiency of tonal information. In the future, it would be important to directly compare the native and non-native patterns within the same models, with a larger corpus. What is also needed is further research to investigate the effect of carry-over coarticulatory information on speech processing and the impact of different L1 tonal and non-tonal backgrounds on the utilization of such cues for speech processing.

5. ACKNOWLEDGMENTS

HNH is currently supported by the Indonesia Endowment Funds for Education (201903223013970). QY is currently supported by the China Scholarship Council (201707110012). YC is currently supported by the NWO Vici grant (vi.c.181.040). We thank Dr. Marjolein Fokkema for sharing her knowledge in GAMM analysis.

6. REFERENCES

[1] Martin, J. G., Bunnell, H. T. 1981. Perception of anticipatory coarticulation effects. *J. Acoust. Soc. Am*. 69(2), 559–67.

[2] Fowler, C. A., Brown, J. M. 2000. Perceptual parsing of acoustic consequences of velum lowering from information for vowels. *Perception & Psychophysics* 62(1), 21–32.

[3] Dahan, D., Tanenhaus, M. K. 2004. Continuous mapping from sound to meaning in spoken-language comprehension: Immediate effects of verb-based thematic constraints. *J Exp Psychol Learn Mem Cogn* 30(2), 498–513.

[4] Beddor, P. S., McGowan, K. B., Boland, J. E., Coetzee, A. W., Brasher, A. 2013. The time course of perception of coarticulation. *J. Acoust. Soc. Am*. 133(4), 2350–66.

[5] Xu, Y. 1997. Contextual tonal variations in Mandarin. *J. Phon.* 25, 61–83.

[6] Shen, X. S. 1990. Tonal coarticulation in Mandarin. *J. Phon.* 18(2), 281–295.

[7] Malins, J. G., Joanisse, M. F. 2010. The roles of tonal and segmental information in Mandarin spoken word recognition: An eyetracking study. *J. Mem. Lang.* 62, 407–420.

[8] Zou, T. 2017. *Production and perception of tones by Dutch learners of Mandarin* [Doctoral dissertation]. Leiden University.

[9] Yang, Q., Chen, Y. 2022. Phonological competition in Mandarin spoken word recognition. *Language, Cognition and Neuroscience* 37, 820–843.

[10] Zou, T., Liu, Y., Zhong, H. 2022. The roles of consonant, rime, and tone in Mandarin spoken word recognition. *Front. Psychol.* 12, 740444.

[11] Shen, J., Deutsch, D., Rayner, K. 2013. On-line perception of Mandarin Tones 2 and 3: Evidence from eye movements. *J. Acoust. Soc. Am.*133(5), 3016–29.

[12] Qin, Z., Tremblay, A., Zhang, J. 2019. Influence of within-category tonal information in the recognition of Mandarin-Chinese words by native and non-native listeners: An eye-tracking study. *J. Phon.* 73, 144–157.

[13] Sun, Y., Green, J., Shih, C. 2021. Anticipatory tonal variations can facilitate spoken-word recognition: A Visual World Eye-tracking study. *Book of Abstracts TAI 2021*, 210–211.

[14] Chen, Y. 2012. Message-related variation. In: Cohn, A., Fourgeron, C., Huffman, M. (eds.), *The Oxford handbook of laboratory phonology*. Oxford University Press, 103–114.

[15] Yuan J.H., Chen Y. (2014), 3rd tone sandhi in Standard Chinese: A corpus approach. *Journal of Chinese Linguistics* 42(1): 218–237.

[16] Magnuson, J. S. 2019. Fixations in the visual world paradigm: where, when, why*?. J Cult Cogn Sci* 3, 113–139.

[17] Qin. Z, Zhang, J. 2022. The use of tonal coarticulation cues in Cantonese spoken word recognition, *JASA Express Letters* 2, 035202.

[18] Cai, Q., Brysbaert, M. 2010. SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PloS ONE* 5(6), e10729.

[19] Sun, C.C., Hendrix, P., Ma, J., Baayen, R. H. 2018. Chinese lexical database (CLD). *Behav Res* 50, 2606–2629.

[20] Pacilly, J. 2022a. CrossSplice.praat. [Praat script]

[21] Pacilly, J. 2022b. ModDuration.praat. [Praat script]

[22] Hallett, P. E. 1986. Eye movements. In: Boff, K. R., Kaufman, L., Thomas, J. P. (eds), *Handbook of perception and human performance* Vol. 1. Wiley Interscience, 10–11.

[23] Wood, S. N. 2020. *mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation.* https://cran.r-project.org/web/ packages= mgcv

[24] van Rij, J., Wieling, M., Baayen, R. H., van Rijn, H. 2022. *itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs*. https://CRAN.R-project.org/package=itsadug

[25] R Development Core Team. 2022. R: A language and environment for statistical computing. https://www.R-project.org/

[26] Porretta, V., Kyröläinen, A.-J., van Rij, J., Järvikivi, J. 2018. Visual world paradigm data: From preprocessing to nonlinear time-course analysis. In: Czarnowski, I., Howlett, R., Jain, L. (eds), *Intelligent Decision Technologies 2017*. Smart Innovation, Systems and Technologies, vol 73. Springer, Cham.

[27] Ito, A., Knoeferle, P. 2022. Analysing data from the psycholinguistic visual-world paradigm: Comparison of different analysis methods. *Behav Res,* 10.3758/s13428-022-01969-3. Advance online publication.

[28] Holm, S. 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* 6(2), 65–70.

[29] Duyck, W., Vanderelst, D., Desmet, T., Hartsuiker, R. J. 2008. The frequency effect in second-language visual word recognition. *Psychonomic Bulletin & Review* 15, 850–855.

[30] Hopp, H. 2017. Individual differences in L2 parsing and lexical representations. *Biling: Lang. Cogn.* 20(4), 689–690.

---

[1] We excluded data from 3 native SC speakers born and grew up in the Southern part of China and 2 learners spoken Hakka or Hokkien in their households.

[2] The difference in the length of learning did not influence learners' eye-fixations patterns.

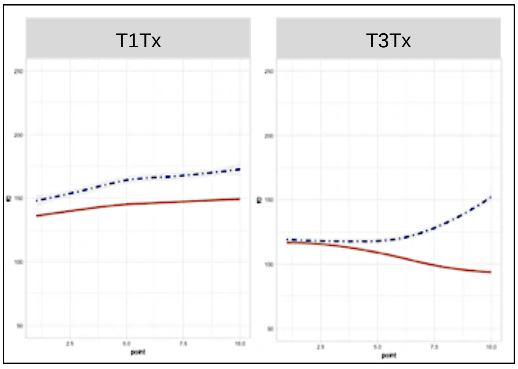
7. APPENDICES

Appendix 1. Stimuli presentation across trial types and example. Number in brackets indicates the original lexical tone of the second syllable before splicing.

Trial type	**Visual Stimuli**		**Auditory Stimuli**	
	Tar	Com	Matched	Mismatched
Critical [1]	宾瓦	宾哇	*bin1*$_{(3)}$*wa3*	*bin1*$_{(1)}$*wa3*
Baseline [2]	宾瓦	垃副	*bin1*$_{(3)}$*wa3*	*bin1*$_{(1)}$*wa3*
Filler 1	宾哇	宾瓦	*bin1*$_{(1)}$*wa1*	*bin1*$_{(3)}$*wa1*
Filler 2	垃副	宾哇	*la1*$_{(4)}$*fu4*	*la1*$_{(2)}$*fu4*
Filler 3	当条	垃副	*dang1*$_{(2)}$*tiao2*	*dang1*$_{(4)}$*tiao2*

Note: [1] The target-competitor contrast is on the 2nd syllable (differ in tone). [2] The contrast lies on both syllables: 1st syllables differ in segments; 2nd syllables differ in segments and tone. The inclusion of baseline trials was aimed at creating a comparable design to that in a four-image VWP.

Appendix 2. Averaged f0 contours of T1 and T3 in target-competitor NW pairs produced by one speaker (n=34 for T1; n=32 for T3).



Tonal contrasts for each panel.
Left: T1T3 (dot-dashed); T1T1 (solid).
Right: T3T3 (dot-dashed); T3T2 (solid).

Appendix 3. Parametric and smooth terms of GAMM analysis of fixations proportion to targets.

		Native listeners				Non-native listeners			
		Est.	SE	*t*	*p*<	Est.	SE	*t*	*p*<
Parametric coefficients	(Intercept)	1.004	0.055	18.307	0.01	0.432	0.054	8.066	0.01
	Cue	-0.242	0.044	-5.495	0.05	-0.109	0.041	-2.694	0.05
	Tone	-0.199	0.044	-4.535	0.05	0.048	0.040	1.183	n.s.
	ToneCue	0.203	0.062	3.263	0.05	0.045	0.057	0.783	n.s.
	Trial	0.398	0.044	9.073	0.01	0.289	0.040	7.151	0.01
	TrialCue	0.185	0.062	2.974	0.05	0.117	0.057	2.044	n.s.
	TrialTone	0.191	0.062	3.062	0.05	0.020	0.057	0.355	n.s.
	TrialToneCue	-0.081	0.088	- 0.917	n.s.	-0.156	0.079	-1.963	n.s.
		edf	Ref.df	F	*p*<	edf	Ref.df	F	*p*<
Approximate significance of smooth terms	s(Time)	6.137	7.072	19.79	0.01	5.494	6.400	5.395	0.01
	s(Time):Cue	6.036	7.299	7.299	0.05	5.921	7.062	1.988	n.s.
	s(Time):Tone	2.269	2.913	2.871	0.05	2.676	3.430	0.912	n.s.
	s(Time):ToneCue	1.000	1.000	1.875	n.s.	3.019	3.880	1.176	n.s.
	s(Time):Trial	7.022	8.108	45.07	0.01	6.855	7.862	4.160	0.01
	s(Time):TrialCue	1.044	1.085	1.532	n.s.	5.541	6.729	2.646	n.s.
	s(Time):TrialTone	1.000	1.001	9.468	0.05	2.671	3.436	0.300	n.s.
	s(Time):TrialToneCue	1.014	1.026	0.416	n.s.	1.001	1.001	1.150	n.s.

We set “critical” as the reference level of Trial, “matched” as the reference level of Cue, and “Tone 3” as the reference level of “Tone”. The first four rows of parametric and smooth terms refer to critical trials, while the last four rows refer to differences in effects (of Time, Cue, and Tone) between critical and baseline trials.