ISCA Archive
http://www.isca-speech.org/archive

4th International Symposium on
Tonal Aspects of Languages (TAL-2014)
Nijmegen, The Netherlands
May 13-16 2014

An acoustic analysis of Mandarin Tone 3 sandhi elicited from an implicit priming experiment

Chierh Cheng [1], *Jenn-Yeu Chen* [1], *Yi Xu* [2]

[1] Department of Chinese as a Second Language, National Taiwan Normal University, Taiwan
[2] Department of Speech, Hearing and Phonetic Sciences, University College London, U.K.

psyjyc@ntnu.edu.tw, chierh.cheng@gmail.com, yi.xu@ucl.ac.uk

Abstract

This study analyses acoustic data elicited from an implicit priming experiment on tone 3 sandhi in Mandarin Chinese (a follow-up to Chen and Chen [1]) with the aim of examining how the tonal contours of lexical T2 and sandhi T3 are realised in a set that shared the initial parts of the word (i.e. the homogeneous condition, e.g. *li³pin³*, *li³xiang³*, *li²dao³*, *li²guo³*) and in a set that did not share the initial parts of the word (i.e. the heterogeneous condition, e.g. *li³pin³*, *yan³jiang³* , *yu²gang³*, *wei²xian³*), respectively. F0 contours and their instantaneous velocity profiles values are used in a functional t-test which allows comparing the mean of two groups of contours along the time domain. We find that in the homogeneous condition, T2 and sandhi T3 are very similar to each other. However, in the heterogeneous condition, the amplitude and speed of the initial falling movements differ, and a time lag in directional change is present between T2 and sandhi T3. This finding suggests that the neutralisation of sandhi T3 and T2 is less complete in the heterogeneous condition than in the homogeneous condition. Additionally, compared to the heterogeneous condition, the longer duration of target syllables found in the homogeneous condition could point to an intricate online articulatory re-adjustment of tonal targets during tonal alternation.

Index Terms: tone sandhi, Mandarin Chinese, implicit priming, functional data analysis

1. Introduction

Production of a group of disyllabic words can be initiated faster when they share the first syllable (homogeneous) than when they do not (heterogeneous). This effect is known as implicit priming of the syllable [2]. Levelt et al. [3] ascribed the implicit priming effect to an encoding scheme which works incrementally from left to right in a suspension-resumption manner that allows for advance preparation of the beginning form of a word.

The concept of implicit priming effect was initially used to explore what unit is activated during word processing and how word preparation works in speech production [2, 4]. Recently, this method has also been used to explore the application of phonological alternation rules, such as tone 3 (T3) sandhi in Mandarin Chinese [1, 5], to examine at which encoding stage of the word-form production system the T3 sandhi rule is applied, and what unit is active during the alternation process.

Mandarin Chinese is a syllabic language with four basic tones (T1: high, T2: rise, T3: low and T4: falling) and one neutral tone (T5: mid). T3 sandhi occurs when two T3 syllables are combined in an appropriate prosodic domain where the first T3 is changed from a low tone to a rising one.

Through four implicit priming experiments, Chen and Chen [1] showed that implicit priming effects were similar for words rendered homogeneous after tone sandhi (Experiment 1: T3T3, T2T3; implicit priming effect: 42 ms) to those that were homogenous prior to tone sandhi (Experiment 2: T3T3, T3T2; implicit priming effect: 49 ms). For such sets, speakers displayed reduced response times indicating an advantage of advance preparation of the first tonal syllable. Reduced response times were also seen when T2 and T3 without tone sandhi were involved (Experiment 3: T2T3, T3T2; implicit priming effect: 51 ms). However, a reduced priming effect was observed for T1 and T3 without tone sandhi (Experiment 4: T1T3, T3T2; implicit priming effect: 27 ms). The authors suggested that these results might indicate the speaker's ability to prepare the first syllable up to the articulatory encoding stage and further suggested that this stage could be added to the production model of Levelt et al. [3].

Owing to these observations, Chen and Chen [1] proposed an articulatory hypothesis to account for the mechanism and locus of tone 3 sandhi during Mandarin Chinese spoken word production. The hypothesis states that tone sandhi is neither a phonological nor a phonetic operation but an articulatory operation which takes place online during tone production. That is to say, speakers may be able to prepare the first syllable up to the articulatory encoding stage when the response words began with T2 and T3 regardless of whether tone sandhi was involved.

To observe the basis of the above articulatory hypothesis, a detailed acoustic analysis is necessary: and this is the purpose of this follow-up study. Here, acoustic experiments were conducted to investigate the tonal contours of T2 and sandhi T3 elicited from an implicit experiment. If, as assumed, a relevant speech production model should consist of an articulatory encoding stage (analogous to smaller articulatory examplars which are populated in speech production phonetic space and are used in calculating speech motor plans, as in Browman and Goldstein [6]), these articulatory gestures will also be utilised in certain planning contexts such as an implicit priming experiment. That is to say, in addition to response time, an advance preparation up to the shared articulatory gesture in the homogeneous condition may well be revealed acoustically.

1.1. Scope of the present study

In the present study, we replicate the first experiment of Chen and Chen [1] with careful attention to the design of the response words and the handling of F0 extraction.

First, we check whether a similar implicit priming effect is obtained (> 40 ms). Confirmation of this effect will to some extent support the findings of Chen and Chen [1] and pave the way for further acoustic analysis.

We then compare F0s from the first syllable of the response for 'T2' (T2T3) and 'sandhi T3' (T3T3), and between the homogeneous and heterogeneous conditions.

This acoustic comparison was carried out not only on F0 contours, but also on their instantaneous velocity profiles (given by the slope of F0 at a single point in time, i.e. its first derivative with respect to time). Additionally, a functional t-test was carried out, which extends the rationale of the well-known t-test to time dependent functions in a principled way. Such an in-depth representation will offer further insight into the articulatory dynamics of actual articulation [7, 8], and therefore help reveal any subtle, yet systematic, patterns of tonal realisation resulting from implicit priming effects.

2. Methodology

2.1. Participants

Ten male native speakers of Taiwan Mandarin, aged between 20 and 30, participated in the study. They all had normal or correct-to-normal vision and had no self-reported speech disorders or professional vocal training. Male speakers were recruited because their formants are easier to track than those of female speakers [9] which are important for the purpose of this study.

2.2. Materials and apparatus

Four sets of word pairs were prepared for the experiment (Table 1). Each pair was made of two two-character words, with the first word serving as a prompt and the second as a response. The two words in a pair shared an obvious semantic relationship, e.g., ju^4-xing^2 ('sentence pattern') and yu^3-fa^3 ('grammar'). For the homogeneous sets, the first characters of the response words in the same set were segmentally identical but their tones were either T2 or T3 and their following tones were all T3s. For the heterogeneous sets, the same words used in the homogeneous sets were rearranged so that the first characters of the response words were no longer segmentally nor tonally identical.

To avoid F0 perturbation effect by initial consonants, semivowels such as glide (e.g. [j, w, y]) or liquid (e.g. [l]) were used as the initials of the first syllables in response. To facilitate segmentation, the initials of the second syllables in response were all obstruents. Only the first syllables of the disyllabic response words were used for acoustic analysis.

An ASUS personal computer (Intel® Core™ i7-3770 CPU at 3.40 GHz) with a 24-inch Benq monitor (32bits, 1920×1080 pixels, 16.67-ms refresh rate) and a Sound-Blaster compatible sound card was used for the experiment. The experiment was programmed in DMDX [10]. The sound card served as the voice key. This was done by setting beforehand a threshold for the sound card. When a vocal response exceeded the threshold, the sound card sent a signal, which stopped the timing routine of the experimental program. The speech signal was received by a condenser microphone connected to the computer. The signal was digitised at 22050Hz in real time using DMDX and stored on the computer's hard disk for later coding and analysis.

2.3. Experimental design and procedure

This experiment followed the design and the procedure of [2] closely. In this experiment, the participants repeated a block of trials three times. Each repetition included a homogeneous condition and a heterogeneous condition, the order of which was counterbalanced across the participants. Within each context condition, the four pairs of words in a set were repeated four times, yielding a total of sixteen trials that constituted a block. The order of the trials was randomized for each set, each context condition, each repetition and each participant. The order of the four sets was also randomized for each context condition, each repetition and each participant. In total, each participant received 384 trials (3 block repetitions × 2 context conditions × 4 sets in a context condition × 4 pairs in a set × 4 set repetitions).

A learning phase preceded the administration of each block of the 16 trials. During the learning phase, the participants were shown an index card on which were printed the four pairs of words to be tested later. The participants learned the association between the two words in each pair until they indicated they had memorized them well. Then, they proceeded to the testing phase. During the testing phase, each trial began with a 1000-Hz warning tone and two short dashed lines flanking a blank space at the center of the screen. The tone and the dashed lines appeared for 200 ms. The prompt word appeared in the previous flanked space 600 ms later. The prompt word stayed on the screen for 150 ms. Another 850 ms elapsed before the trial ended. The participants were told to say out loud the target word upon seeing the prompt word, as quickly and accurately as possible. If no response was initiated within 1000 ms of the presentation of the prompt word, a feedback tone of 500 Hz was sounded for 200 ms. The next trail began after another 200 ms. If a response was initiated, then the sound was recorded for 1000 ms. The next trail began after another 200 ms.

The participants were seated 60 cm from the screen. Their eye level was at the center of the screen. The microphone was adjusted to an appropriate position to ensure that the voice key could be activated effectively.

2.4. Measurements

Sound files were segmented by the first author, a native Taiwan Mandarin speaker. To delimit the first syllables in the disyllabic response words, points nearest the zero crossing at both edges of the first syllables were used. Onset times of the first syllables were used to represent response times. All ten participants' response times were used for measuring the implicit priming effect (Sec. 3.1). Among them, five out of the ten participants were randomly selected for detailed acoustic analysis (Sec. 3.2).

Extraction of F0 from the first syllables in the disyllabic response words was first carried out with the vocal cycle marking of the Praat program [11] and then with manual repair of octave jumps and other distinct irregularities using a Praat script [12]. Each target curve was sampled at ten equidistant points. F0 values were converted into semitones and the mean of the ten samples subtracted from each contour to help reduce speaker variability. Functional data analysis was then applied to the processed data as describe below.

2.5. Functional data analysis and functional t-test

Functional Data Analysis (FDA) refers to a set of tools that extend ordinary multivariate statistics to the domain of functions. In this study, the FDA framework is used to compare F0 contours that have different durations in a principled way. An FDA session requires that the input curves are represented as functions, in our case functions of time $f(t)$ (hence the term 'functional' in FDA). Standard smooth interpolation techniques were customarily applied at this stage. In particular, we applied B-splines-based smoothing with roughness penalty [13].

A crucial aspect of the functional representation of data lies in the fact that the same time axis must be used in all curves. This means that all curves are normalised to have the same duration. Since the original tone dyads do not have the same duration, a proper normalization has to be carried out. In our case, the normalization was based on the sampling scheme described in 2.4 (i.e. ten equidistant points within every target syllable). The total duration is congenitally indicated as 1 (*not* the average duration in seconds).

Comparisons between pairs of F0 contours sets were carried out as follows. An average function was extracted from each set of curves, i.e. a function whose value at every point in time is the average of all the functions in the set at that same time (e.g. Fig.2a). The same goes for F0 velocity profiles, which were calculated before the time normalization, so as to preserve the original velocity values. Functional t-test was also applied to the groups of functions whose averages were displayed (e.g. Fig.2b). Functional t-test compares the mean functions of two groups of functions sharing the same time axis and determines *where* those mean curves differ, at a certain confidence level. Operationally, a battery of ordinary t-tests is carried out in the time dimension. The confidence value for the test statistics were obtained by resampling techniques (see [13], for further details). All FDA operations were carried out using the R package 'fda' [14, 15].

3. Results and discussion

3.1. Error rate and response time (RT)

Each participant's responses were categorised into: correct responses, incorrect responses, no responses, hesitations, and corrected responses. Only correct responses were considered for analysis. All other responses were considered errors. The overall error rate was about 6%. The error rate in the homogeneous condition (3%) was lower than that in the heterogeneous condition (9%).

For response time (RT) analysis, tokens with response times outside the threshold of mean response time ± 3s.d. were removed as outliers (0.3%). Fig.1 shows the distribution of response times of trimmed data with interaction lines by context and sandhi conditions.

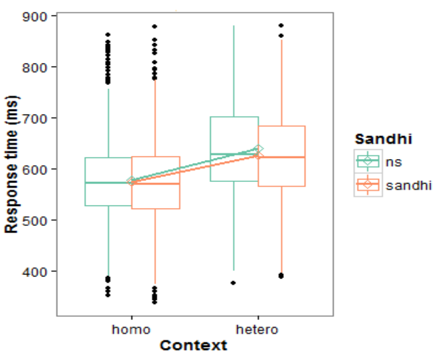

Figure 1: *Distribution of response times (in milliseconds) and interaction lines by context and sandhi conditions. The horizontal line in the middle of a box plot displays the median and the diamond symbol marks the mean.*

We analysed the data of response time by using a linear mixed effects model. We entered context and sandhi (with interaction) in the model as fixed effects. As random effects, we had intercepts for subjects and items, as well as by-subject and by-item random slopes for the fixed effects. P-values were obtained by likelihood ratio tests of the full model with the effect in question against the model without the effect in question. All data was analysed using R [14] and lmerTEST [16].

The analysis indicated that context affected response time ($\chi^2(5)=493.7$, $p<.0001$), and resulted in an increased response time in the heterogeneous condition by 58.1ms±6.8 (standard errors). The sandhi effect did not affect response time ($p=.31$) and the context effect did not differ between sandhi condition, i.e. no interaction was found ($p=.19$).

These results suggest that context does have an effect on response time. Specifically, our results suggest that the speakers responded quicker when the response words are in the homogeneous condition than in the heterogeneous condition, independent of the sandhi condition. These results seem to be in line with the previous findings of Chen and Chen [1].

3.2. Acoustic results

3.2.1. F0 contours and velocity profiles

In this section, we report the results from comparing F0 of the first syllable of the participants' response between T2 and sandhi T3 in the homogeneous condition (Fig.2), and between T2 and sandhi T3 in the heterogeneous condition (Fig.3).

Figures 2 and 3 both consist of four plots: (a) the normalized average F0 contours of T2 and sandhi T3 as a function of time (note that the mean F0 has been subtracted from the overall F0s so as to remove some of the inter-speaker variability, as mentioned in section 2.4), (b) the corresponding functional t-test on F0 contours between these two curves, (c) the normalized average velocity of T2 and sandhi T3 as a function of time and (d) the corresponding functional t-test on velocity between these two curves. In t-test plots (b) and (d), both blue lines represent the 0.05 critical values for the t-statistic, the higher flat one being a more conservative threshold. The solid red line represents the time dependent observed statistic.

In Fig.2, i.e. the comparison between T2(T3) and T3(T3) in the homogeneous condition, both curves appear to be very similar to each other in terms of both F0 contours and their velocity profiles. The general trend between these two curves is an initial fall (negative velocity) for the two-fifth of the total time course, followed by a rise (positive velocity) beginning from 0.4 normalised time units until the end of the syllables. In other words, for both T2 and sandhi T3, the rise does not begin until at least 40% of the syllable is completed.

The similar F0 amplitudes of T2 and sandhi T3 shown in Fig.2a were supported by the t-values in Fig.2b being less than 2 along the whole syllable. Note that a marginal significance also occurs at the 0.4 normalised times, which might indicate a slightly lower position for sandhi T3 to approach than for T2 before they both switched to rise. But overall these two curves were very similar to each other. Fig.2d also indicates no significant differences between these two curves in terms of the speeds of changing direction.

Turning to the comparison between T2(T3) and T3(T3) in the heterogeneous condition (shown in Fig.3), we can see that both curves are qualitatively similar to each other, which is akin to their homogenous counterparts, and share the same falling-rising movements (cf. Fig.2ac and Fig.3ac). However, unlike their homogeneous counterparts, the quantitative differences between T2 and sandhi T3 are evident in at least three respects.

First, Fig.3a shows that the overall amplitude of falling movements is larger in sandhi T3 (0.9 semitones, onset minus minimum) than T2 (0.5 semitones), which was apparent seeing a higher onset position and deeper valley of sandhi T3 than of T2. Second, the two separate zero crossings in Fig.3c indicate that T2 changes direction from a fall to a rise earlier than sandhi T3. Third, in addition to T2's earlier change of direction from fall to rise, its initial falling speed was slower (as indicated by a smaller absolute velocity value, st/s) than that of sandhi T3. The slower initial falling movement of T2 also indicates a reduced falling speed of T2 in the heterogeneous condition (Fig.3c) compared to that of T2 in the homogeneous condition (Fig.2c).

Overall, sandhi T3-T2 neutralisation is evident in both the homogeneous and heterogeneous condition, with the latter being less complete. The less similar initial tonal movements and the existence of a time lag of direction change between T2 and sandhi T3 in the heterogeneous condition might suggest that the quantity used in calculating speech motor plans differ somewhat per context condition.

For the homogeneous condition, the close resemblance of T2 and sandhi T3 might indicate a motor priming effect stemmed from the shared tonal properties, i.e. the same initial falling gesture of T2 and sandhi T3. Although being less significant in the homogeneous than in the heterogeneous condition, the lower F0 minimum of sandhi T3 than T2 might be indicative of tonal alternation from an underlying low target of T3 to a rising target of sandhi T3.

Furthermore, the heterogeneous condition might involve a new retrieval of the underlying mono-tonal target of T3, which makes sandhi T3 target closer to its canonical one. In other words, speakers may have given the underlying target of T3 more weight than its sandhi target in the heterogeneous condition.

3.2.2. *Duration*

Four two-sample t-tests were conducted to compare the durations of the first syllables of responses by context and by sandhi condition (Table 2). The durations of T2 and sandhi T3 in the homogeneous condition were both significantly longer than in the heterogeneous condition. No significant differences were found between the durations of T2 and sandhi T3 within the same context condition.

Table 2. *T-tests on mean durations (in milliseconds) of the first syllables in disyllabic response words by context and sandhi conditions.*

	Homo	Hetero	Significance
T2(T3)	164.5*	156.6	$p<.01$
T3(T3)	161.9*	154.5	$p<.05$
Significance	$p=.39$	$p=.43$	

The reasons why the durations of T2 and sandhi T3 differ in different context conditions is not clear. The implicit priming effect is about 58 ms (Sec. 3.1) and the difference in the duration of the target syllables between the two context conditions is around 8 ms. The implicit priming effect found in the homogeneous condition does not seem to retain its advantage after the initiation of articulation. In other words, an earlier initiation of the articulation does not seem to lead to a faster implementation (or completion) of actual articulation. One potential explanation could be that in the homogeneous condition the speakers might have used the 8 ms to make online articulatory adjustment for tonal alternation. Unlike the heterogeneous condition (for which articulatory gestures are presumably less comfortably prepared), in the homogeneous condition the initial articulatory gestures of T2 and T3 were purposely kept amenable for change and therefore extra time might be needed for readjusting the ensuing gestures.

Surely, there could be another line of reasoning. For example, a longer duration of the target syllables found in the homogeneous condition could be a feature of tonal stability, which can be an indication of the completeness of sandhi T3-T2 neutralisation in a certain planning context. Notwithstanding our attempt to interpret the current results, why initially primed articulation led to a longer implementation of target syllables in the homogeneous condition is not clear and we welcome thoughtful comments from readers.

4. Conclusions

In this paper, an acoustic comparison was performed to explore sandhi data elicited from an implicit priming experiment. Results from F0 and velocity values indicated a lower degree of neutralisation for sandhi T3 and T2 in the heterogeneous condition than in the homogenous condition. This acoustic evidence may support a potential priming effect extending to the articulatory encoding stage in a certain planning context. However, it is unclear why the primed syllables in the homogeneous condition led to a longer articulatory implementation. It could be indicative of an intricate online sandhi mechanism in calculating speech motor plans for the advanced preparation of articulatory gestures in a potential sandhi domain. Motivated by this exploratory analysis, we believe it is worth re-examining the motor planning account of tonal variation. Specifically, how an articulatory encoding stage, as proposed by Chen and Chen's [1] articulatory hypothesis, is related to the gradient realization of tone sandhi. This question remains to be deciphered and evidence of this kind should be helpful in furthering our understanding of tonal changes in general.

5. Acknowledgements

This research is partially supported by the "Aim for the Top University Project" of National Taiwan Normal University (NTNU), sponsored by the Ministry of Education, Taiwan, R.O.C. and the "International Research-Intensive Center of Excellence Program" of NTNU and National Science Council, Taiwan, R.O.C. under Grant no. NSC 103-2911-I-003-301.

6. References

[1] Chen, J.-Y., and Chen, T.-M., "Mechanism and locus of Tone 3 sandhi during Mandarin Chinese spoken word production," Journal of Chinese Linguist., accepted in 2013.

[2] Meyer, A. S., "The time course of phonological encoding in language production: The encoding of successive syllables of a word", Journal Memory and Language, 29(5):524-545, 1990.

[3] Levelt, W.J.M., Roelofs, A., and Myers, A.S., "A theory of lexical access in speech production", Behavioral and Brain Sciences, 22:1-75, 1999.

[4] Chen, J.-Y., Chen, T.-M., and Dell, G.S. "Word-from encoding in Mandarin Chinese as assessed by the implicit priming task," Journal of Memory and Languages, 46:751-781, 2002.

[5] Politzer-Ahles, S., and Zhang, J. "The role of tone sandhi in speech production: Evidence for phonological parsing," in Proc. Tonal Aspects of Languages (TAL-03), 2012.

[6] Browman, C.P, and Goldstein, L. "Gestural specification using dynamically-defined articulatory structures," Journal of Phonetics, 18:299-320, 1990.

[7] Cheng, C., Chen, J.-Y., and Gubian, M., "Are Mandarin Sandhi Tone and Tone 2 the Same or Different? The Results of Functional Data Analysis", in Proc. Pacific Asia Conference in Language, Information, and Computation (PACLIC-27), Taipei, Taiwan, 2013.

[8] Gauthier, B., Shi, R., and Xu, Y., "Learning phonetic categories by tracking movements", Cognition, 103:80-106, 2007.

[9] Fulop, S. A., Speech Spectrum Analysis. New York; Springer, 2011.

[10] Forster, K. I., and Forster, J. C., "DMDX: A windows display program with millisecond accuracy", Behavior Research Methods, Instruments, & Computers, 35:116-124, 2003.

[11] Boersma, P., and Weenink, D. "Praat: doing phonetics by computer [Computer program], Version 5.3.52". Online: http://www.praat.org/retrieved, last viewed on 12 June 2013.

[12] Xu, Y. 2013., "ProsodyPro—A Tool for Large-scale Systematic Prosody Analysis", in Proc. Tools and Resources for the

Analysis of Speech Prosody (TRASP 2013), Aix-en-Provence, France, 2013.
[13] Ramsay, J. O. and Silverman, B. W., “Functional Data Analysis—2nd Ed”, Springer, 2005.
[14] R Core Team, “R: A language and environment for statistical computing”, R Foundation for Statistical Computing, Vienna, Austria. Online: http://www.R-project.org/, last viewed on 16 December, 2013.
[15] Ramsay, J. O., Hadley Wickham, Spencer Graves and Giles Hooker, “fda: Functional Data Analysis. R package version 2.3.8”. Online: http://CRAN.R-project.org/package=fda, 2013.
[16] Alexandra, K., Bruun Brockhoff, P., and Haubo Bojesen Christensen, R., “lmerTest: Tests for random and fixed effects for linear mixed models (lmer objects of lme4 package). R package version 2.0-3”. Online: http://CRAN.R-project.org/package_lmerTest, last viewed on 16 December, 2013.

Table 1. *Stimuli sets. Prompt words corresponding to each response are shown in parentheses. Homo: Homogeneous; Hetero: Heterogeneous.*

	Homo 1	**Homo 2**	**Homo 3**	**Homo 4**
Hetero 1	(賄賂)禮品 $li^{3}pin^{3}$	(草稿)演講 $yan^{3}jiang^{3}$	(出海)魚港 $yu^{2}gang^{3}$	(安全)危險 $wei^{2}xian^{3}$
Hetero 2	(現實)理想 $li^{3}xiang^{3}$	(眉毛)眼角 $yan^{3}jiao^{3}$	(低能)愚蠢 $yu^{2}chun^{3}$	(巨大)微小 $wei^{2}xiao^{3}$
Hetero 3	(當兵)離島 $li^{2}dao^{3}$	(漁業)沿海 $yan^{2}hai^{3}$	(滋潤)雨水 $yu^{3}shui^{3}$	(植物)葦草 $wei^{3}cao^{3}$
Hetero 4	(水份)梨果 $li^{2}guo^{3}$	(逾時)延緩 $yan^{2}huan^{3}$	(句型)語法 $yu^{3}fa^{3}$	(訂金)尾款 $wei^{3}kuan^{3}$

(a) homo. F0 contours *(b) t-test on F0 contours* *(c) homo. velocity profiles* *(d) t-test on velocity profiles*

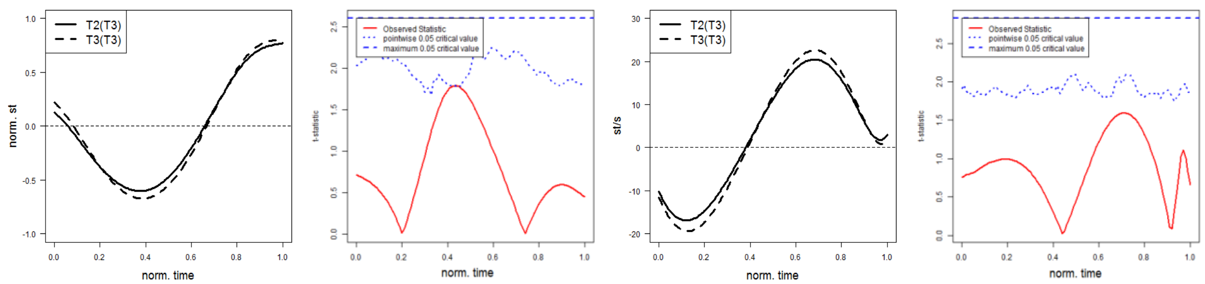

Figure 2: *T2(T3) and T3(T3) in the homogeneous condition (black). (a) Normalised average F0 contours of T2 (solid) and sandhi T3 (dashed). (b) Functional t-test on normalised average of F0 contours. (c) Normalised average F0 velocity profiles and (d) Functional t-test on normalised average F0 velocity contours.*

(a) hetero. F0 contours *(b) t-test on F0 contours* *(c) hetero. velocity profiles* *(d) t-test on velocity profiles*

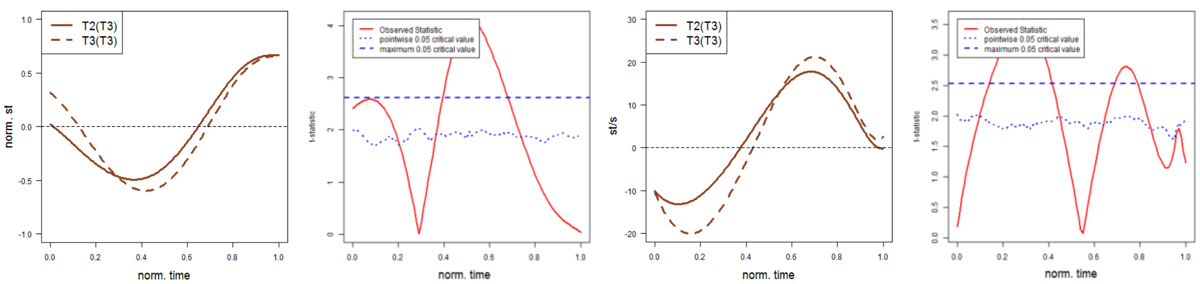

Figure 3: *T2(T3) and T3(T3) in the heterogeneous condition (brown). (a) Normalised average F0 contours of T2 (solid) and sandhi T3 (dashed). (b) Functional t-test on normalised average of F0 contours. (c) Normalised average F0 velocity profiles and (d) Functional t-test on normalised average F0 velocity contours.*