

Exploring the Syntax-phonology Interface: The Effect of Freestanding Form

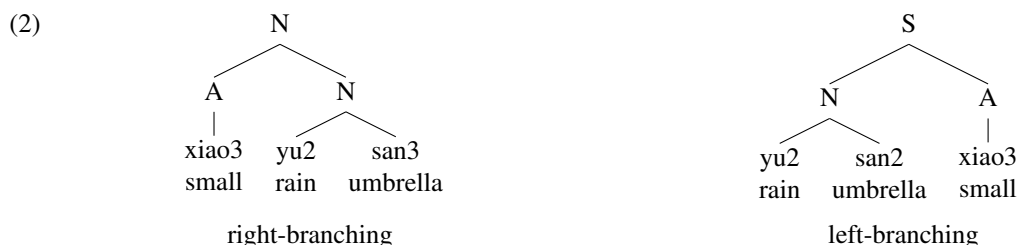
Yujing Huang
Harvard University

1 Introduction

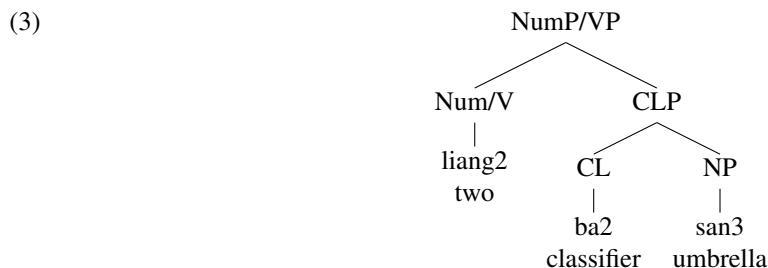
Tone 3 sandhi in Mandarin is a phenomenon where a contour tone Tone 3 changes to a rising tone Tone 2 when it precedes another Tone 3 (see (1)).

(1) 214 → 35/___214 (or Tone 3 → Tone 2/___ Tone 3)¹

It has been claimed that this sandhi is conditioned by syntax and applies cyclically (Chen 2000, Duanmu 2007, a.o.). For example, the right-branching phrase *xiao yu san* ‘small umbrella’ has the sandhi pattern *xiao3 yu2 san3* and the left-branching phrase *yu san xiao* ‘the umbrella is small’ has the sandhi pattern *yu2 san2 xiao3*.



The previous proposals run into problems with a structure that is characterized by a numeral/verb followed by a classifier and a noun (i.e. numeral/verb + classifier + noun). In this type of phrase, the structure is transitionally analyzed by syntacticians as (3), but the sandhi pattern is documented as 223. Previous research explained the sandhi pattern by a cliticization rule (see Poteet 1985). However, I will show that this cliticization rule cannot explain the intuition of Mandarin native speakers. Section 2 will review the cliticization rule. Section 3 will show an experiment of Mandarin native speakers’ judgment on the sandhi pattern of numeral + classifier + noun phrases. Section 4 will provide a proposal to explain the judgment data. Section 5 will conclude the paper.



* I thank Kevin Ryan, C.-T. James Huang, Isabelle Charnavel, Michael Kenstowicz, Lauren Eby Clemens and the audience at Annual Meeting on Phonology 2015 for their feedback. This work is funded by the Mind Brain Behavior Interfaculty Initiative at Harvard.

¹ The numbers are from Chao (1980) and represent the height of the pitch, where 1 is the lowest and 5 is the highest. The space between 1 and 5 is divided evenly. Therefore, Tone 2 rises from middle to high. Tone 3 falls from half low to low and then rises from low to half high.

2 Previous studies

Compared to the abundant literature and various theories on more canonical Mandarin Tone 3 sandhi (e.g. (2)), the sandhi pattern of numeral/verb classifier noun (henceforth XCN) phrases is relatively overlooked and the existing explanations are fairly similar.

The first proposal is advocated by (Poteet, 1985) as a cliticization rule. In his example, he used a phrase of verb + classifier + noun: *mai ba san* (to buy an umbrella). In his account, the classifier cliticizes onto the verb, and the verb and the classifier together form a prosodic word. This account has been adopted by (Shih, 1986) and later modified by (Chen, 2000). In Chen (2000), the cliticization happens as a restructuring rule and restructures the VP. It is not clear whether the restructuring is intended to happen at the morpho-syntactic level or prosodic level. No matter which is the intended level, the cyclic sandhi domain will become $((mai\ ba)_1\ san)_2$. Therefore, the sandhi pattern should be the same as a left-branching tree.

However, native speakers seem to have different intuitions on the sandhi patterns of a left-branching tree and a XCN phrase. This will be shown in the next section.

3 Experiments

Study 1: Judgment on Tone 3 sandhi patterns

The Tone 3 sandhi patterns have been documented in the literature. However, the judgments are from the authors with limited examples. In addition, most papers overlook the possible variations in some structures. Therefore, we conducted a study to test the sandhi patterns with naive native speakers.

Method

Participants As different dialects differ in their tonal inventory, this study only recruited speakers from Beijing. This is because the Beijing dialect has the same tone as Mandarin so we exclude the possible Tone 3 sandhi pattern variation due to dialectal influence. Twenty-four native Mandarin speakers from Beijing were recruited.

Materials The materials consist of 36 3-syllable chains with three different syntactic structures. They are: left-branching structure (i.e. $[[\sigma\sigma]\sigma]$) such as *yu san xiao*, right-branching (i.e. $[\sigma[\sigma\sigma]]$) such as *xiao yu san* and XCN structure such as *mai ba san* (see Fig. 1). The stimuli vary in two dimensions. The first dimension is the syntactic structure, i.e. left-branching (LB), right-branching (RB) and XCN. The second dimension is the sandhi pattern. As the previous studies on T3S have observed (cf. Duanmu (2007)), T3S variation is generally available in RB but not LB. This study therefore included the sandhi pattern structure as a variable. Within each structure, there are two tone patterns, i.e. 223 and 323, and thus, the study contains six conditions: LB 223, LB 323, RB 223, RB 323 XCN 223 and XCN 323.

$[[\sigma\sigma]\sigma]$	$[\sigma[\sigma\sigma]]$	XCN
<i>[[yu san] xiao]</i>	<i>[xiao [yu san]]</i>	<i>[mai [ba san]]</i>
umbrella is small	small umbrella	buy umbrella

Fig. 1

Each condition has half 323 and half 223 tone patterns. The audio stimuli are read by the author who is a monolingual native Mandarin speaker. Stimuli are read at the normal speech rate in a sound-proof phonetic booth.

Procedure The study took place online. Stimuli were presented to the participants in a random order. Because the sandhi Tone 3 is the same as lexical Tone 2 in perception (see Peng 1996), when the participant hears the tone pattern 223, it is possible that the first tone is an underlying Tone 2. To avoid potential lexical ambiguity, in each trial, the participants saw the Chinese characters of the phrase they heard. On each page, the participant saw one phrase and an audio button to play the sound file. The participants were asked to rate the pronunciation of the sound of the corresponding characters on a 1 to 7 scale based on how acceptable the tones are pronounced, where 1 is the unacceptable score and 7 is totally acceptable. After participants

provided the judgment, they proceeded to another stimulus with the same experimental structure described above.

Results

A linear mixed effect model (with structure, sandhi pattern as fixed effect, and lexical item, participant as random intercept) shows a main effect of structure and sandhi pattern, and an interaction between structure and sandhi pattern (all $p < 0.01$). The interaction shows that the acceptability of the sandhi pattern varies across different syntactic structures. The differences in the acceptability patterns are not due to differences in lexical items.

To interpret the above results in more details, the means for each condition are computed. The mean for LB 223 is 6.76 and LB 323 is 2.20. The mean for RB 223 is 3.80 and RB 323 is 6.59. As the results for verb + CL + N and Numeral + CL + N does not differ ($p = 0.80$ for the 223 pattern and $p = 0.34$ for the 323 pattern), the two groups are collapsed for the analysis. The mean XCN 223 is 5.80 and the mean for XCN 323 is 4.22. The distribution is shown in Fig. 2.

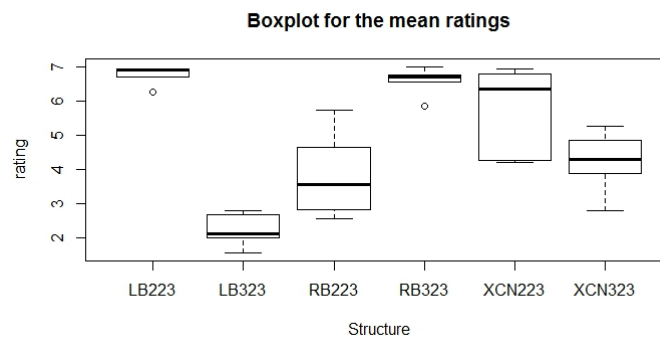


Fig. 2

After comparing the means, we found that the LB 223 is not different from RB 323 or XCN 223 (Mann-Whitney-Wilcoxon test $p = 0.37$ and $p = 0.14$ respectively). All of these forms are well-formed. RB 223 and LB 323 are significantly different (Mann-Whitney-Wilcoxon test $p = 0.005$). RB 223 and RB 323 are also marginally different (Mann-Whitney-Wilcoxon test $p = 0.049$). The difference between XCN 223 and XCN 323 is not significant but marginal ($p = 0.06$). If the item number or participant number were increased, the result would likely be significant.

Critically, XCN 323 is different from LB 323 (Mann-Whitney-Wilcoxon test $p = 0.006$). Because previous literature assumes the same prosodic structure between the two forms, this difference is unexpected. As we can see from the results, XCN patterns differently from both the left-branching structure and the right-branching structure. If it were a left-branching structure, we would expect XCN 323 to be unacceptable while it is significantly different the unacceptable form LB 323. If it were a right-branching structure, we would expect XCN 323 to be more acceptable than XCN 223 which is clearly not the case.

Discussion

The above results show that the sandhi pattern of XCN does not pattern with either a left-branching structure or a right-branching structure. Therefore, the cliticization rule which assumes the same prosodic structure of LB and XCN cannot explain the sandhi pattern of the XCN structure. It also shows that RB structure and XCN structure both allow variation while LB structures do not. It is consistent with the judgment in previous studies. The two variants for RB structures (i.e. RB 323 and RB 223) are not equal and the two variants of XCN also tend to be different. Therefore, the previous accounts in the literature are challenged because they cannot explain the judgment pattern described above.

As XCN differs from both RB and LB, the remaining logical possibility is a flat syntactic structure $[\sigma\sigma\sigma]$. Empirically, XCN does differ from a RB structure because CL + Noun is not a free-standing constituent. Therefore, it is possible that the prosodic structure of XCN is the same as the flat structure which lacks syntactic inner constituents.

Study 2: Prosody of a flat structure

To test the hypothesis in the discussion section above, the default prosodic structure of a flat syntactic structure needs to be investigated. This section describes a study testing the prosodic boundaries of a three-syllable chain with no inner structure. The rationale is: after controlling for the syllable structure, the only difference between syllable boundaries should be caused by prosodic grouping. As it is not feasible to control for syllable structure with XCN phrases, a classic example of flat structure, namely number chains, is used instead.

Method

Participants 5 Mandarin native speakers with an average age of 25 and a minimal education level of college were recruited. None of them speaks any other dialects. Therefore, dialectal influences are excluded.

Materials To control for syllable structure, the same number is repeated in test sentences to make sure that the syllable structures are identical. For ease of acoustic analysis, the number 8 is used as a non-Tone 3 stimulus because it is the only number in Mandarin that has a stop as the onset and no nasal coda. The number 9 is used as the Tone 3 stimulus because it is the only number in Mandarin that has an affricate as the onset and no nasal coda². To make the situation more natural, the numbers were provided as phone numbers, zip codes or car numbers with the frame sentence “Name’s X is Y”.

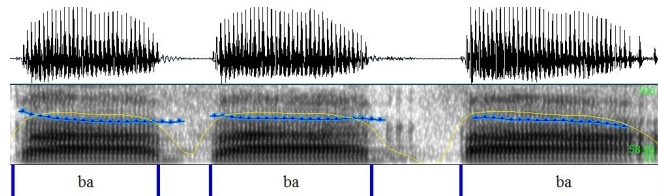
Procedure After a pilot study, eight sentences with hypothetical number chains were chosen and presented (e.g. John’s phone number is 221-462-7529.) Six of them had the same number repeating 3 times (e.g. Mary’s number is 6713-157-999.) Four of the sentences had repeating numbers with Tone 3 (the number 9 in Chinese).

Participants were asked to read the sentences they were given naturally at their normal speed. As number chains are in general hard to read, the participants were given chances to practice as much as they needed and asked to record themselves once they were fluent with the numbers. In this way, speech errors and unexpected pauses are excluded.

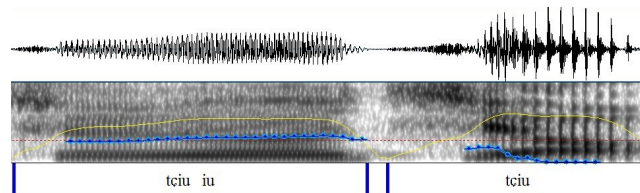
Acoustic Analysis The recordings were analyzed with Praat (Boersma & Weenink, 2013). The duration of the silence between the end of the vowel and the beginning of the onset of the next syllable were measured and documented.

It is found that the duration of the second silence is in general greater than the first one (see (4)). In some cases, some speakers reduced the first silence to zero and even omitted the onset of the second syllable (see (5)), but only one participant in one instance omitted both and never was there any case where they omitted the second but not the first.

(4) 8-8-8



(5) 9-9-9



² No number with a Tone 3 has a stop onset.

Results and discussion

A paired t-test was conducted to compare the duration of the silence; it showed that the two pauses are significantly different ($t=-3.73$, $p<0.001$). Since the order of comparison is Pause 1 vs. Pause 2 and the confidence interval is $(-0.022, -0.0066)$, we can conclude that Pause 2 is greater than Pause 1. The distribution of the difference between Pause 1 and Pause 2 (Pause 2- Pause 1) is shown in Figure 3.

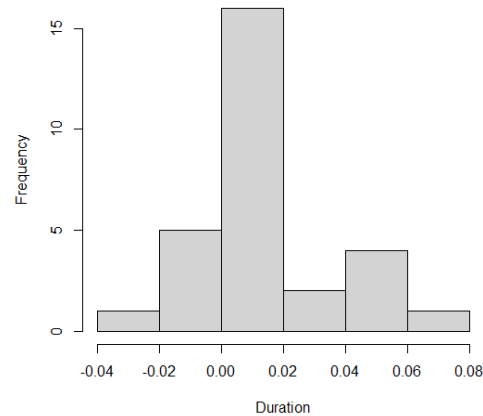


Fig. 3

The difference in duration indicates that there are two layers of prosodic structure. Therefore, the prosodic structure of flat structure 3-syllable Tone 3 chains can be seen as $((\sigma\sigma)\sigma)$. This structure is a LB one. Therefore, we would expect 223 to be the best candidate. However, since this is a prosodic, not a syntactic, structure, it may be different from a syntactic LB structure, and as has been shown above, the acceptability variation in this group is different from that of an LB group.

4 Towards an analysis

Study 1 shows that XCN differs from both a left-branching and right-branching structure in its Tone 3 sandhi pattern. I previously observed that the inner constituent of XCN is not a free-standing structure. In this section, I will propose an account to explain the sandhi pattern of XCN. However, before we arrive at such an analysis, we need to have an account for the Tone 3 sandhi pattern in general.

The first question to ask is what Tone 3 sandhi is. This paper follows Yip (2002) and attributes Tone 3 sandhi to the Obligatory Contour Principle (OCP):

OCP-3: no adjacent Tone 3s are allowed.

This constraint as defined can account for Tone 3 sandhi while keeping other tones from changing. Following Duanmu's (2007) observation and Chen's (2000) insight, OCP-3 is an across-the-board constraint which can apply across any syntactic domain.

As OCP-3 is a markedness constraint, there should be a corresponding faithfulness constraint Ident-IO(Tone) ranked lower than it:

Ident-IO(tone): assign a violation to every tone that has been changed.

There are two possibilities that can account for the restricting effect of syntax in Tone 3 sandhi. The first is that Tone 3 sandhi applies cyclically in each domain which is defined by Match theory (Selkirk, 2011) or Edge correspondence such as Generalized Alignment (McCarthy & Prince, 1993). However, Lin (2005) shows that such an analysis will result in the same amount of constraint violation for LB 323 and RB 223 which are judged to be different in the perceptual test. Therefore, this paper proposes a different account.

In this paper, an output-output (OO) correspondence constraint is adopted to capture the cyclic effect. Because the OO constraint maximizes the "phonological identity between morphologically related output forms" (Kager, 2004), it also bridges between phonology/prosody and morphology/syntax.

Ident-OO(tone): the tones need to be identical in derivationally related outputs.

The following tableaux show that with the above constraints, the basic pattern with RB and LB structures can be explained.

In (6), the LB structure $[[yu\ san] xiao]$ has a free-standing inner constituent $yu\ san$ which can serve as the base. Therefore, Candidate (a) violates Ident-OO once because the second tone is different from the base. Candidate (b) violates Ident-OO twice because both of the first two tones are different from the base. Candidate (c) does not violate Ident-OO but it violates OCP-3 which is ranked higher. Therefore, the winning candidate is Candidate (a) as in (6).

(6)

Input: $[[yu3san3]xiao3]$		OCP	OO	Ident-IO
Base: $yu2san3$				
a.	$\text{ㄩㄥˊ ㄙㄢˊ} ((22)3)$		*	* *
b.	$(32)3$		* *!	*
c.	$((23)3)$	*!		*

In (7), the RB structure $[xiao [yu\ san]]$ has a free-standing inner constituent $yu\ san$. No candidate violates Ident-OO in the tableau. Candidate (c) violates OCP-3 and Candidate (b) violates Ident-IO one more time than Candidate (a). Therefore, the winning candidate is Candidate (a) as in (7).

(7)

Input: $[xiao3[yu3\ san3]]$		OCP	OO	Ident-IO
Base: $yu2san3$				
a.	$\text{ㄩㄥˊ ㄙㄢˊ} (3(23))$			*
b.	$(2(23))$			* *!
c.	$(2(33))$	*!	*	*

Comparing (6) and (7) it can be seen that RB 223 violates IO one more time than RB 323 while LB 323 violates OO one more time than LB 223. If we assume that a violation of a higher ranked constraint (OO in this case) is worse than violating a lower ranked one (IO in this case), we can explain why LB 323 is less acceptable than RB 223.

There is a remaining group: the XCN group. Because XCN is a right-branching structure in syntax. If we assume that is also its prosodic structure and follow the above analysis, we will make a wrong prediction about the sandhi pattern as shown in (8).

(8)

Input: $[3[33]]$		OCP	OO	Ident-IO
a.	$\text{ㄩㄥˊ ㄙㄢˊ} (3(23))$			*
b.	$\text{ㄩㄥˊ ㄙㄢˊ} (2(23))$			* *!
c.	$(2(33))$	*!		*

Candidate (b) should be the winner while Candidate (a) wins in this tableau. Since there is no violation before IO for either Candidate (a) or (b), switching the order of the constraints will not help.

As has been mentioned earlier, the inner constituent of XCN is not a free-standing form. It is possible that because the inner constituent is not a free-standing structure, this constituent is transparent to phonological evaluation. Therefore, the Number + Classifier + Noun is treated as a whole and the T3S patten is accounted for by prosody.

It was shown in Study 2 that a 3-syllable flat structure has prosodic structure $((\sigma\sigma)\sigma)$. This can be explained by a module of OT constraints.

First, the experimental data as well as the previous studies show that the prosodic structure is binary. Here, this paper adopt Chen's (2000) Binary constraint:

Binary: the prosodic structures are binary.

To avoid introducing the concept of foot, whose definition is not clear in Chinese, alignment constraints are used instead.

However, traditional alignment constraints do not differentiate XCN and other RB structures. To integrate the effect of free-standing form, the alignment constraints need to be defined as below.

Align-L: given that Φ is a prosodic phrase and XP is the relevant free-standing syntactic domain, Align (Φ , L, XP, L).

Align-R: given that Φ is a prosodic phrase and XP is the relevant free-standing syntactic domain, Align (Φ , R, XP, R).

As Study 2 showed, when there is no inner constituent, the default prosodic structure is $((\sigma\sigma)\sigma)$. Therefore, Align-L needs to rank above Align-R³, as tableau (9) shows.

(9)

Input: [333]	Binary	Align-L	Align-R
a. (333)	*!		
b. $\left[\left[\begin{array}{c} \text{3} \\ \text{3} \end{array} \right] \right] \text{3}$			*
c. (3(33))		*!	

Putting it all together, the Tone 3 domains are the product of interaction between syntax and prosody. The tableaux (10) and (11) show how the correct candidates are chosen in LB and RB structures respectively.

(10)

Input: [[yu3san3]xiao3]	OCP	OO	Binary	Align-L	Align-R	IO
Base: [yu2san3]						
a. (3(23))		* *!		*		*
b. ((32)3)		* *!				*
c. (323)		* *!	*			*
d. $\left[\left[\begin{array}{c} \text{2} \\ \text{2} \end{array} \right] \right] \text{3}$		*				* *
e. (2(23))		*		*!		* *
f. (223)		*	*!			* *

(11)

Input: [xiao3[yu3san3]]	OCP	OO	Binary	Align-L	Align-R	IO
Base: [yu2san3]						
a. $\left[\begin{array}{c} \text{3} \\ \text{3} \end{array} \right] \text{3}$						*
b. ((32)3)					*!	*
c. (323)			*!			*
d. (2(23))						* *!
e. ((22)3)					*!	* *
f. (223)			*!			* *

The tableau of XCN is illustrated as (12)

³ The brackets specify syntactic boundaries.

(12)

Input: [333]	OCP	OO	Binary	Align-L	Align-R	IO
a.  ((32)3)					*	*
b. (3(23))				*!		*
c. (323)			*!			*
d. (2(23))				*!		**
e.  ((22)3)					*	**!
f. (223)			*!			**


However, the analysis above does not quite work for XCN yet as shown in (12). The correct candidate should be Candidate (e) while Candidate (a) is chosen. To choose the right candidate, two more constraints need to be added. The first constraint deals with the observation that Tone 3 sandhi is sensitive to boundary strength. When the boundary between two Tone 3s is weak, Tone 3 sandhi is more likely to happen. The second constraint deals with the observation that Tone 3 is never fully realized in non-final positions with no boundaries. Even when it precedes tones other than Tone 3, it undergoes some changes.⁴

Ident.IO(T)_prepausal: Tone 3 remains faithful in prepausal positions.


*T3]₀: T3 is marked when there is no boundary between it and the following tone.

Therefore, (10) becomes (13), (11) becomes (14) and (12) becomes (15).


(13)

Input: [[yu3san3]xiao3]	OCP	OO	Binary	Align-L	Align-R	Ident.IO(T)_prepausal	*T3] ₀	IO
a. ((32)3)		**!					*	*
b. (3(23))		**!		*				*
c. (323)		**!	*				*	*
d. (2(23))		*		*!		*		**
e.  ((22)3)		*				*		**
f. (223)		*	*!					**

(14)

Input: [xiao3]yu3san3]	OCP	OO	Binary	Align-L	Align-R	Ident.IO(T)_prepausal	*T3] ₀	IO
a. ((32)3)					*!	*	*	*
b.  (3(23))								*
c. (323)			*!				*	*
d. (2(23))						*!		**
e. ((22)3)					*!	*		**
f. (223)			*!					**

(15)

Input: [333]	OCP	OO	Binary	Align-L	Align-R	Ident.IO(T)_prepausal	*T3] ₀	IO
a. ((32)3)					*	*	*!	*
b. (3(23))				*!				*
c. (323)			*!				*	*
d. (2(23))				*!		*		**
e.  ((22)3)					*	*		**
f. (223)			*!					**

With this ranking, we are able to explain why XCN differs from both RB and LB structures.

What is important here is that the mapping between syntax and prosody is sensitive to whether the intermediate output in the morpho-syntactic derivation is free-standing; that is, the alignment constraints are sensitive to whether the constituent is a free-standing form or not.

⁴ The pitch contour of Tone 3 in the terminology of Chao (1980) is 214 but it is realized as 21 when it is followed by any other tone.

5 Summary

This paper examines a case where not being a free-standing form changes the sandhi pattern. The proposal is similar to Base-Identity proposed in Kenstowicz (1996), where the base needs to be an independent word. The difference between the current proposal and Base-Identity is that Base-Identity predicts what will happen when there is an independent word morphologically related to the form under evaluation, while the current proposal predicts what will happen when there is no independent base. Base-Identity is realized by comparing two independent morphologically related forms, while the current proposal associates two forms by redefining the alignment of boundaries. Finally, this paper shows that the Mandarin Tone 3 sandhi, which was believed to be a cyclic application, can instead be derived with an output-output constraint.

References

- Boersma, Paul & David Weenink (2013). Praat: doing phonetics by computer [computer program], version 5.3. 39.
- Chao, YR (1980). A system of tone letters. *Le MatVtre PhoneTique* .
- Chen, Matthew Y (2000). *Tone sandhi: Patterns across Chinese dialects*, vol. 92. Cambridge University Press.
- Duanmu, San (2007). *The phonology of standard Chinese*. Oxford University Press, USA.
- Kager, René (2004). *Optimality theory*. Cambridge University Press.
- Kenstowicz, Michael (1996). Base-identity and uniform exponence: alternatives to cyclicity. *Current trends in phonology: Models and methods* 1, 363–393.
- Lin, Hui-shan (2005). Prosodic correspondence in tone sandhi. *UST Working Papers in Linguistics* 1, 229–265.
- McCarthy, John J & Alan Prince (1993). Generalized alignment. *Yearbook of morphology* 79–153.
- Peng, Shu-hui (1996). *Phonetic implementation and perception of place coarticulation and tone sandhi*. Ph.D. thesis, Ohio State University.
- Poteet, Stephen (1985). Tone sandhi and the phonological structure of mandarin. *Ms, UCSD* .
- Selkirk, Elisabeth (2011). The syntax-phonology interface. *John Goldsmith, Jason Riggle and Alan CL Yu (eds.), The handbook of phonological theory* 435–484.
- Shih, Chi-Lin (1986). *The prosodic domain of tone sandhi in Chinese*. Ph.D. thesis, University of California, San Diego.
- Yip, Moira (2002). *Tone*. Cambridge University Press.